THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XVI | September, 1925 | Number 6

THE NATURE OF INTELLIGENCE

ROBERT MORRIS OGDEN

Cornell University

In 1921, the *Journal of Educational Psychology* published a symposium in which 14 American psychologists stated what they individually conceived "intelligence" to be.[1] Since then Professors Charles Spearman of London and L. L. Thurstone, one of the contributors to the Symposium, have each published books on the subject.[2]

It is the purpose of this paper to comment briefly upon the definitions of intelligence offered by these 15 psychologists, and to suggest yet another definition for consideration.

Broadly speaking, I find that 7 of the 14 contributors to the Symposium agree fairly well in their definition of intelligence as *learning capacity*. Two use this term explicitly; while 5 speak of "good responses," "adaptation," "capacity for knowledge," "acquiring-capacity," and learning in terms of "conditioned responses;" all of which, it seems to me, are very close to the general concept of learning.

Despite this agreement, I am disturbed by the absence of definite criteria for intelligence. No doubt the measure of intelligence is directed upon the "good responses" and the "successful acquisitions"; but when is a response "good," and what is a "successful" achievement? These 7 writers tell us little more than that the good is the true and factual, or that success is something desirable.

If we turn to the other 7 writers in the Symposium, we find less agreement upon the definition of intelligence, but likewise little that is of substantial aid in understanding what intelligence is. Two contributors venture no answer at all; one because he thinks the data too

[1] Intelligence and Its Measurement: A Symposium. Vol. 12, pp. 123–147, 195–216, 271–275.

[2] "The Nature of 'Intelligence' and the Principles of Cognition." London, 1923. "The Nature of Intelligence." London and New York, 1924.

meager, the other for lack of interest in the question. Two others content themselves with empirical or traditional statements of mental capacities; one warning us that in the tests now current too little heed is being given a considerable list of mental factors which he then proceeds to name. Of the three that remain, one calls intelligence "a biological mechanism by which the effects of a complexity of stimuli are brought together and given a somewhat unified effect in behavior"; but concerning this "somewhat unified effect" we learn nothing in particular.

A simple, though to me at least a suggestive, statement refers intelligence to "abstract thinking," and lauds the verbal tests as being more appropriate than any others in determining this little understood process of mind. Finally, Thurstone's answer, which I mention by name since the author has elaborated it into a book, is the only one that departs both from the empirical level and from the usage of traditional though undefined terms of psychology. Thurstone begins with instinct. "Mental life," he writes, "is an irreversible process beginning with the life-impulse and terminating in the successful overt act. The stimulus may be thought of as a means for specifying the approximate act which is mental."[1] In other words, the springs of action are instinctive desires, the stimuli being the modifiers which induce consciousness. Intelligence is "(*a*) the capacity to inhibit an instinctive adjustment; (*b*) the capacity to re-define the inhibited instinctive adjustment in the light of imaginally experienced trial and error; (*c*) the volitional capacity to realize the modified instinctive adjustment into overt behavior to the advantage of the individual as a social animal."[2] In his book he writes: "The development of intelligence is measurable by the incompleteness of the motive at which it can become focal. If the motive becomes focal as an imaginal experience before the stimulus has even appeared, the animal has in consequence of this capacity an increased control in the satisfaction of its wants with minimum risk." And again: "Intelligence, considered as a mental trait, is the capacity to make impulses focal at their early, unfinished, stage of formation."[3]

Thurstone's definition has, at least, the merit of being a genetic approach to the problem; furthermore, it implies that intelligence has to do with inhibitions which must be overcome "largely in rendering

[1] "The Nature of Intelligence," p. 23.

[2] "Symposium," pp. 201–2.

[3] "The Nature of Intelligence," pp. 79 and 159.

conscious an unfinished act."[1] This reference of intelligence to the conscious completion of an act is, I think, highly important. I am, however, less content with a re-definition of the act "in the light of imaginally experienced trial and error," because in my opinion trial and error do not adequately describe the process of intelligent discovery. As for the end attained by intelligence "to the advantage of the individual as a social animal," this end, I fear, suggests the mysteries of vitalistic teleology.

What, indeed, is lacking in all these definitions of intelligence is a scientific definition of *ends*, which will clarify the process of intelligent response, and serve as a useful principle in devising and experimenting with mental tests. We should all admit, I suppose, that tests are empirical devices the validity of which is determined by statistical correlations with other empirical data, such as school grades, teachers' estimates, and various performances and attainments of the individual tested; but I can not understand how any definition of intelligence derived therefrom can do more than summarize in very general and usually very vague terms an empirical procedure which has been followed and found useful in testing.[2] A true definition of intelligence must be given in scientific terms—terms which themselves can be defined and the data they represent experimented upon. A definition of this order is calculated to explain why the tests are useful as measures of intelligence, and when and how they fail. It would also lead to experiment with less empirical and more narrowly defined methods of measurement.

Such an attempt has been made by Professor Spearman, who is well known to students of mental measurement for his theory of a general and purely quantitative factor of intelligence to which specific and variable factors are added. According to Spearman the *rationale* of testing is found in the effect obtained from pooling specific tests in order that the general factor which underlies them may thus have an opportunity to express itself. When, however, Spearman comes to describe the psychology of intelligence we find him leaving empirical grounds which, at least, have the virtue of actuality, and taking refuge in logic.

---

[1] "Symposium," p. 204.

[2] Thus J. L. Stockton in his Definition of Intelligence, *Psychological Review Monograph Series*, 137, 1921, ascribes intelligence to an organism "when it possesses the ability to influence its destiny through the utilization of an inner, active, non-predictable selective factor which chooses on the basis of similarity." This factor appears to be otherwise known as *judgment*.

I shall describe his explanation briefly because it has many points of contact with an explanation which I myself have employed in my textbook of psychology,[1] in other writings,[2] and also during many years in my courses.[3]

Professor Spearman finds three principles involved in intelligence: (1) Granted sentience, "any lived experience tends to evoke immediately a knowing of its characters and experiencer."[4] The form which experience thus takes implies "some sort or other of what may be called 'material,'" but whereas the "form" is "only the aggregate of those attributes of the occurrence which remain in all instances similar," the "'material' is composed of such attributes as may vary from one instance to another."[5] Here, apparently, the general factor is the "form" and the special factors "the material."

Spearman calls his second principle the eduction of relations: "The mentally presenting of any two or more characters (simple or complex) tends to evoke immediately a knowing of relation between them."[6] As a corollary it is stated that a relation once apprehended may thereafter act as a fundament in a new relation.

Spearman's third principle is the eduction of correlates: "The presenting of any character together with any relation tends to evoke immediately a knowing of the correlative character."[7]

An act of intelligence consists, therefore, in evoking relations between fundaments given as known, thereby creating new fundaments out of these evoked relations; it extends also to the evocation of a fundament as yet unknown which satisfies the conditions of a known relation given with a single equally known fundament.

I have used a very similar scheme in my textbook and in my teaching. While I never regarded the relation as being directly existential, in the way Spearman seems to do, I have attempted to show that a relation, once it *subsists* between two existential fundaments, becomes a subject for abstraction. Since one relation can be compared with another, a "notion," as I called it, can in this way be

---

[1] An Introduction to General Psychology." New York, 1914.

[2] *Cf.* especially Mental Activity and Conscious Content in "Philosophical Essays in Honor of James Edwin Creighton." New York, 1917, pp. 290–306.

[3] A more detailed discussion of Professor Spearman's book will be found in *The American Journal of Psychology*, vol. 36, 1925, pp. 140 ff.

[4] *Op. cit.*, p. 48.

[5] *Op. cit.*, p. 49.

[6] *Op. cit.*, p. 63.

[7] *Op. cit.*, p. 91.

derived of the relationship itself. This derivation of a notion from two or more pairs of related terms was intended to explain the origin of what were called imageless thought-contents. My scheme also embraced Spearman's third principle of evocation by which an at-first-unknown fundament is sought and found when the consciousness of one fundament to which a relation is attached indicates the direction and determines the character of a second unknown fundament which fulfills the conditions of the relation.

As I now regard this scheme it appears somewhat artificial, and its grounds more logical than psychological; yet I believe it can be amended in a way that will afford a clearer insight into the actual processes of cognition, and the growth and differentiation of intelligence.

Regarding any datum of observation, any [illegible]ied cause of behavior, in its genetic aspect, its first appearance can be likened to the emergence of an articulate figure from an inarticulate ground. We can indicate that which is "lived" or "undergone"—as Spearman calls it—by A, and the less articulate ground from which it emerges by not-A. At this lowly level, apprehension is something more than a static presentation in sensory terms, since it is also a dynamic process with a temporal course. Obviously, no analysis of so vague a figure can be undertaken either in terms of sensory or motor "elements," for the simple reason that no such elements are as yet defined by the response. We can, however, assume for it those qualities of urgency which are felt in moments of desire, and which are registered in the direction of an action which fulfills such a desire. In other words, the response has a certain form of sentience and feeling, and of behavior, which sets it off from other possibilities, not realized at the time.

The next step in the procedure of cognition, or intelligent response, is the definition of this form, and this definition refers to what we may term the *contour* of the figure. We are applying the term *figure* both to the apprehension of the experience in the usual psychological sense and also to the form of the behavior as behavior is understood in an "objective" psychology. This usage would, of course, be nonsense if we meant by figure an aggregate of sensory and affective processes on the one hand, and an aggregate of physiological processes on the other hand. The term *figure* is chosen because it suggests form rather than content; the assumption being that the figure of a square object, for instance, and the apprehended squareness may be identical figures, although from an artificially analytic point of view which deals

separately with the physical nature of the square object and the psychological nature of the apprehended squareness, they would be worlds apart. Upon the emergence of a figure the next step in cognition is the definition of the figure's contour; a process which refers both to the behavior of response and to the apprehension of the figure, because the one depends upon the other.

How does this definition of contour take place? Our answer must be given in terms of the conditions that prevail, and if we would understand how to control behavior, these conditions must be studied separately and analytically. A figure of intelligence can not be analyzed. It simply *is* or *is not.* But as conditions vary, so the figure of response varies. In a certain sense every figure of behavior, or of apprehension, is appropriate to its conditions; nevertheless, the combination and selection of conditions, which is a part of every response, give rise to figures which evince greater and lesser degrees of articulateness of the response made by the total psycho-physical organism to its environment. An act of intelligence is an act which reflects these higher degrees of definiteness, wherein the figure becomes sharply drawn and highly articulate.

If we lived in a chaotic world, in which chance alone reigned supreme, there would be no standards of excellence. But that is not the case. All experience is shot through with æsthetic moments in the contemplation of what, at the time, is unique in its perfection; so, too, all behavior has its moments of perfect adaptation to the conditions under which it is aroused. These are the complete and perfect figures, the contours of which are clearly and effectively defined. Whenever these moments appear, they stand out with a peculiar insistence which enables them to serve as patterns for all future responses. While every definition is not of necessity *right* or *true,* it is nevertheless in terms of what we have called articulateness that the investigation of rightness and truth are carried on. The definition of contour as here understood is jointly conditioned by the structure of the perceived object, by the maturation and previously established state of readiness of the sensorium, and by the general attitude of the entire psycho-physical organism at the time of apprehension.

What a definition of contour *does* to a figure that was not done in its first stage of vague and ill-defined emergence is attributable to the internal articulateness of the configuration. The terms of such an intelligent configuration are not merely A and not-A, but include also the *dimensions* of A; these dimensions being the specific characters of

A which supply the basis of all comparison and all specific differentiation. I should therefore say that the relation which Spearman educes is a feature of the configuration, and also that his eduction of correlates is the completion of a figure by supplementation of a missing member.

The essential difference between Spearman's interpretation of cognition and the varieties of intelligence, and the one here suggested for consideration may be summarized as follows: Spearman employs the conception of an aggregate of materials and forms which he treats on the same level as experienced data. These materials and forms can then be supplemented and enriched with new materials and new forms by the act of eduction. Our conception denies the validity of the aggregate as a basic principle of experience, and regards cognition and intelligence as processes involving the constructive membership implied by a figural whole. The apprehended unit is from the start an organized whole, although it is at the same time one which may be capable of further organization through a definition of its contour and inner articulation. This definition, however, is accomplished neither by the associative principle of accretion, nor by an alien and somewhat mysterious act of eduction, but solely by an integrative principle of organic membership which is inherent in the factual data themselves. Integration thus promotes the selection and retention of those figures of apprehension and response that are intrinsically most appropriate and most definite with reference to the entire psycho-physical complex.

When Spearman remarks upon the surprising validity of the Ebbinghaus Completion Test as a measure of intelligence, he at the same time emphasizes the fact that intelligence is a striving for completeness—an originally conditioned desire to bridge all gaps which otherwise render experience incomprehensible, and behavior ineffective. Indeed, comprehension and effectiveness are themselves but terms descriptive of a membered figure, the contour of which is complete and definite because its internal structure is complete and definite.

What then is the nature of intelligence? I answer that the nature of intelligence is revealed by the contour and articulateness of experience. Whether the mode of the experience in question is a picture to the eye, or a melodic sequence to the ear; whether it is a rhythmically balanced response felt as a course of action, or as a train of thought, the process in each case is intelligent in the degree in which it is articulate, compact, sharply outlined, and therefore precise and effective. In order that we may understand the wide variety of intelligences of

which the human mind is capable, we must study these and many other configurations of apprehension and response; we must determine their norms of completeness under the various conditions of maturation and special training which call them forth, and their bearings one upon the other at the empirical level of concrete behavior where tests of intelligence are necessarily made. A "good" response is good because it conforms to a norm of complete behavior, and it is at the same time *true* and *factual* because truth and fact are both measured in terms of norms which the experience of behavior dictates. The discovery of truth or fact is determined by the conditions under which the event occurs as well as by the definition of the apprehended content. Intelligence, measured by sharpness of contour, must therefore be determined with reference to the conditions under which the test is made, and not merely in terms of the phenomena observed and reported by the individual tested.

From this point of view a true empiricism replaces the inexact descriptions of concrete behavior which the tests have been wont to employ. The pattern of any experience or response can be observed and described, but we must be careful that our terms are truly descriptive so as to avoid, on the one hand, artificial and remote implications of formal logic and, on the other hand, the vague generalities of common-sense. Intelligence, I maintain, is always something which can be described as a figure or pattern of consciousness and behavior, appearing completely and uniquely under the conditions appropriate to its arousal. A good test of intelligence is a set of conditions under which various individuals produce or fail to produce the clearly defined pattern which the conditions have warranted. Thus, the criterion of a good single test is that it affords an unequivocal answer.

It is, however, possible to vary the conditions in such a way that a graded series of tests, like those of Köhler's[1] detour-board, will give us a measure of intelligence in terms of degrees of difficulty. A similar result is afforded by a "battery" of tests, with this difference, however: In a simple situation such as Köhler's detour where the animal must first push a desired object *away* from itself before the object can be drawn inwards and laid hold of, the pattern of the behavior is readily comprehended, whereas a "battery" of tests can only be treated statistically, since we do not know wherein the difficulties lie or what precisely constitutes the figures of intelligence upon which the graded results depend.

[1] Köhler, W.: "The Mentality of Apes." New York, 1925, pp. 239 ff.

In order that we may gain a truer scientific insight, which will enable us to devise more exact instruments of measurement, we must approach the problem of intelligence in terms of definite responses, the contours of which can be discerned and described in the same way in which the behavior of Köhler's apes, when confronted with a desired object placed on the detour-board, can be discerned and described. For this reason I am inclined to believe that more attention should now be given in our psychological laboratories to an experimental study of the tests themselves, and of the materials which they employ.

LENGTH OF SENTENCES AS A CRITERION OF A CHILD'S PROGRESS IN SPEECH

MARGARET MORSE NICE

Amherst, Mass.

A child's skill in sentence structure affords a simple criterion of his advancement in the mastery of speech. I wish to propose a division of speech development into four stages dependent on the sentence; *viz.*, single words; early sentences; the short sentence of three to four words and the complete sentence of six to eight words. My reasons are two: Simplicity of determination and the real significance of each stage. The mere recording of a series of 30 or more sentences and finding their average length give us a cross-section of a child's speech development and enable us to know many things as to his progress in this distinctly human achievement.

The single word and early sentence stages have long been recognized, but the importance of the short and complete sentence seems to have been unnoticed. Unfortunately I did not discover their significance myself until two years ago, when all of my children were using complete sentences, therefore my data are not as full as I would have wished.

The Single Word Stage

Practically all children begin to talk with single words. This stage has been reported with 18 English-speaking children, as lasting from 4 to 9 months the average duration amounting to 6 months. In the case of 17 children of other nationalities, the use of single words lasted from 4 to 12 months, the average period being 6.6 months.

The Early Sentences

The combining of two words previously used singly marks a distinct advance in the child's linguistic journey. The age at which the first sentence appeared with 20 English-speaking children varied from 13 to 27 months, the average being 17.5 months. With 20 children of other nationalities, the first sentence was uttered at ages varying from 13 to 23 months, the average being 18.3 months.

When a child first begins to put words together, a large part of his talking is still done in single words. For instance, during the all-day conversation of my fourth daughter, H, at two years (her first sentence had appeared four months earlier) 7 sentences of three words were used, 352 of two words, while 683 words were employed singly. This

gave an average sentence length of 1.35 words. The speech of Kirkpatrick's daughter at 28 months was evidently more advanced than that of H for she averaged 2.35 words to a sentence in a series of 100 sentences. Single sentences in this period may be fairly long, but a representative series will give an average of more than one and less than three words.

The majority of these early sentences are incomplete; this was true of 96 of the 100 sentences of Kirkpatrick's daughter. They consist mainly of nouns, verbs, adverbs, and adjectives; while pronouns, prepositions, and conjunctions appear seldom or never.

Children differ a good deal as to the number of words they can use when they first combine two; with 19 children the vocabularies ranged all the way from 14 words to 180. All of these vocabularies consist mainly of nouns; throughout this early sentence stage the child's quota of words exhibits a "baby ratio," that is, 65 per cent or more of nouns, a high proportion of interjections and sometimes also of adverbs.

How long does it take for a child to progress from his first sentence of two words to the point where he is expressing himself in sentences that average three and four words? My data on this matter are meager. With my second daughter, D, the early sentence stage lasted about 3 months, with H about 7 months, and with my third daughter, R, at least 13. Perhaps from 4 to 7 months is the usual range.

The Short Sentence Stage

The short sentence stage is very characteristic of the little child; it is definitely beyond the early sentence period but it is decidedly distinct from the later complete sentence stage. The child who is using this type of sentence has not yet mastered inflections and is omitting many minor words.

The following table gives the average length of sentences of 16 English-speaking children at ages varying from two to ten years.[1]

[1] All but one set of these averages were based on consecutive series of sentences: Eleven on 33 to 63 sentences (Miller, H at 2½, 3½ and 4, Owen, Howard, D at 2½ and 3½ L, and N) 6 on 100 sentences (Kirkpatrick, A and E at 3 and 10); five on series ranging from 130 to 217 (Moore, Helliwell, Trettien, D at 4 and 8); one on 344 (R at 5); another on 455 (H at 3) and 6 on all-day conversations (Brandenburg, H at 2, D at 3, R at 4 and E at 5). Boyd collected 1240 sentences of his child when 20 to 24 months old and 1250 when she was 3 and 4 years old. The fact that H, E and D do not show consistent increase is due to the shortness of some of the samples; in every one of these cases, except H at 2½ and 3, the greater the number of sentences, the smaller was their average length, within the limits, of course, of their stage.

TABLE I.—AVERAGE LENGTH OF SENTENCES AT DIFFERENT AGES

Name	Child	Age in years											
		2	2⅙	2⅓	2½	3	3½	4	5	6¾	7	8	10
Moore	Boy	4.1											
Helliwell[1]	Girl	3.6											
Miller	Boy	....	5.2										
Nice, H	Girl	1.35	...	....	3.7	4.4	7.2	6.1					
Owen[1]	Boy	....	...	....	...	...	4.5						
Howard[1]	Boy	....	...	....	...	...	3.5						
Kirkpatrick	Girl	....	...	2.35	...	4.0	...	7.0					
Boyd	Girl	3.4	...	....	...	6.2	...	7.0					
Trettien	Girl	....	...	....	5.0								
Brandenburg	Girl	....	...	....	...	6.6	...	7.5					
Nice, E	Girl	....	...	....	...	7.1	...	...	6.1	...	...	...	8.5
Nice, D	Girl	....	...	....	4.4	3.8	7.1	6.4	...	...	...	6.9	
Nice, R	Girl	....	...	....	...	...	...	3.9	6.2				
L[1]	Girl	....	...	....	...	...	...	...	4.1				
N[1]	Boy	....	...	....	...	...	...	...	3.1				
A[1]	Girl	....	...	....	...	...	...	...	...	5.0	6.0		

[1] Mrs. Sue V. Helliwell of Milwaukee, Wis., kindly gave me a series of her daughter's sentences. Owen and Howard were two little neighbors whom I had the opportunity of observing. L and N are twins, and A is their older sister.

Two of these children, H at 2 years and Kirkpatrick's daughter at 28 months were in the early sentence stage. All but three of the 31 other instances fall in one or another of two distinct groups; in 13 cases the sentences averaged from 3.5 to 4.5 words and in 15 they averaged from 6 to 8 words.

The 3.5 to 4.5 word[1] sentence first appears in the table at two years, it is common at three, and occasionally persists as late as four or even five years. At two years, three of four children were in this stage, the other not having yet attained it; at three years, three of six children averaged this length of sentence. It is evident that rather advanced

[1] Further evidence that the short sentence seems to fluctuate between 3.5 and 4.5 words and does not go as low as 3 words nor as high as 5, was shown in the all day conversation of my daughter R (Nice, '20), where the sentences are averaged hour by hour throughout the day. The hourly averages were: 3.5, 3.6, 3.6, 3.7, 3.7, 3.8, 4, 4, 4, 4, 4, 4.5.

children are using this type of sentence at two years—and some doubtless earlier—whereas very many have not yet reached this point in speech development. Many three-year-olds are in this stage; besides the children cited in the table, I have recorded sample sentences of five other children of this age; two of whom used the short sentence and three the complete sentence of six to seven words; this gives five three-year-olds in the short sentence stage and six in the complete sentence stage. At three and a half there are examples of two children in each period. Probably the majority of four-year-olds have passed beyond this phase; my only example of the contrary is that of R who was very loath to learn to talk like other people. The twins who were using short sentences as late as five years of age, although of normal intelligence, were markedly retarded in their speech.

The duration of this stage doubtless varies considerably, but we have little information on the subject. With D, R and H it seems to have lasted somewhat over a year; some children must pass through it more quickly while others probably linger longer in it.

One reason for the shortness of this average sentence is the simplicity of the little child's ideas. This is indicated by the rarity of compound and complex sentences. Of 50 sentences of the five-year-old L only two were compound and none complex; of 50 of her brother N, one was complex and none compound. Kirkpatrick's daughter at three used 5 compound and 11 complex sentences in a series of sentences averaging 4 words, while at four in a similar series averaging 7 words there were 29 compound and 14 complex. A much greater contrast between the two stages was shown by R, for of 350 sentences averaging 3.9 words at four years, only 3 were compound and none complex; while of 50 complete sentences a year later, 14 were compound and the same number complex.

The other reason for the shortness of these sentences lies in the fact of the omission of many minor words and the resulting incompleteness of a large proportion of the sentences. The words most frequently left out are parts of the verb "to be," prepositions, conjunctions, auxiliary verbs, articles and pronouns.

As to the proportion of incomplete sentences, of the 100 averaging 4 words each of Kirkpatrick's daughter at 34 months, 62 were incomplete; 38 lacking the subject, and 44 the assertive verb. With my children such sentences formed the following proportions (in series of 43, 100, 200 and 218 consecutive sentences respectively): H at 30 months, 54 per cent; at 3 years, 37 per cent; D at 3 years, 21 per cent

and R at 4 years, 30 per cent. Of the 50 sentences of L at 5 years, 40 per cent were incomplete, with N, 34 per cent.

The smallest known vocabulary with which the 3 to 4 word sentence is associated is that of Moore's son of 475 words, while the largest is R's of 1135 words. As R exhibited such an atypical history of speech development, this relationship may be exceptional with her, *i.e.*, she added words to her vocabulary faster than she developed the ability to use them properly. It may be that a more usual upper limit of size of vocabulary associated with this type of sentence is denoted by H's and D's vocabularies at three years—804 to 856 words; however, some children probably have smaller vocabularies still.

As to the character of the vocabulary correlated with the short sentence, the "stable ratio" in regard to the proportions of the parts of speech seems to have been reached, *i.e.*, when nouns comprise from 50 to 60 per cent of the words, and verbs from 20 to 24 per cent (the vocabularies having been recorded without regular inflections of verbs). This stable ratio is remarkably uniform with children—English-speaking, French and German[1]—after they have acquired a certain size of vocabulary —usually between 400 and 600 words, but sometimes less and occasionally more. (Almost without exception the small vocabulary of the baby not well advanced in speech shows a very high proportion of nouns and a small quota of verbs.) That these proportions once attained remain about the same is shown in all cases where vocabularies of the same child have been taken on succeeding years beginning with or later than his third birthday.[2]

My evidence for the coincidence of the attainment of the short sentence and the stable ratio of the vocabulary is based on the records of my daughters D, R and H. This was true also of Moore's son at two years. In no other case are both sets of data available.

To sum up then, the main characteristic of the short sentence stage is the fact that the child is not inflecting his verbs and that he is omitting less essential words. The short sentence exhibits an excess of nouns, and a lack of articles, auxiliary and copulative verbs, preposi-

[1] This ratio is shown in the four examples of vocabularies of French children from 33 months to 7 years of age given by Descoeudres (1921: 195), in those of a French and a German girl of two years cited by the Sterns (1907: 61–62) and in all the vocabularies of the Scupin's son—a German boy—from 2 to 6 years.

[2] The vocabularies of Boyd's (1914) and Brandenburg's (1915, 1916) daughters and of my daughter D were recorded at 3 and 4 years, of Pelsma's daughter at 2, 3 and 4 years, of the Rowe's son at 4 and 6, of my daughter E at 3, 4, 5 and 6, and of the Scupin's son at 2, 3, 4, 5 and 6.

tions and conjunctions. From 20 to 60 per cent of these sentences are incomplete. As to the child's vocabulary at this stage it probably contains from 400 to 800 or 1000 words, 50 to 60 per cent of which are nouns and 20 to 24 per cent verbs.

The Transition Stage

There is, of course, a transition stage between the short and complete sentence, but it must be of short duration for I never happened to record it with any of my children, while I have countless examples of the other stages. I have been able to find only three instances of what seemed to be a transition stage; in each of these the average number of words was five. No details are given as to the character of the talk of Trettien's daughter at the age of 2½. In the 50 sentences of the Miller's son at 26 months, all words are properly inflected and only one sentence is incomplete ("are" was omitted). Perhaps the reason that this child's sentences were shorter than the typical complete sentence was that, on account of his being so very young, his ideas were still simple enough to be adequately expressed in less than 6 to 7 words.

The last case is that of my niece A, at 6¾ years, who, although of normal intelligence and in the Second Grade at school was greatly retarded in her speech. At this time she had not mastered inflections, and 32 per cent of her sentences were incomplete. Thus we have two very different pictures in these five-word sentences—one of a precocious and the other of a retarded child. It may well be that neither of these children presented a good example of an average transition stage.

The Complete Sentence

The complete sentence of 6 to 8 words is characteristic of the speech of the more advanced children of three years of age. The earliest date for the attainment of this stage that I happen to have is 33 months (my eldest daughter E), but it doubtless appears earlier with precocious children. All the series of sentences of children with normal speech development from four years on, show about the same sentence length. Even in the case of an adult's conversation (partly with children and partly with other adults) the average of 144 sentences was 8 words.

The attainment of this length of sentence—as the *average* of a series, not the exception—means that inflections have been practically mastered, and that the majority, if not all of the sentences are complete. D and H completed their mastery of inflections at the age of 3½ and at the same time their sentences showed this average length, whereas six months before the average had been from 3 to 4 words. The same thing was true of R at 4½. Of course inflections themselves add little to the sentence length—in fact the added "wills" and "woulds" hardly make up for the lost "dids"—but their attainment seems to be fairly simultaneous with the consistent use of minor words formerly omitted, *i.e.*, if a child has reached the point of saying "comes" and "ate" he will not as a rule omit "is" and "I," "the" and "a." Besides this the greater complexity of the sentences now used makes for increased length. A comparison of the proportions of compound and complex sentences in short and complete sentences has already been made.

As to completeness, in a child who had just attained this stage—A at 7 years—23 of the 100 sentences averaging 6 words were incomplete. Kirkpatrick's daughter's 100 sentences at 46 months averaging 7 words also showed 23 incomplete, in 9 of which the subject was omitted and 14 the assertive verb. These cases would seem to be rather exceptional, however. Every one of the hundred sentences of my daughter E at 33 months was complete.

The smallest known vocabulary with which the 6 to 8 word average sentence is associated is about 1000 words—that of E at three. D and H must have had about the same number of words when they attained it.

The complete sentence stage, then, is characteristic of the child who has practically reached the level of simple adult speech; many children of three years have attained it and most have by four years. It is associated with the mastery of inflections and as far as we know at present by a size of vocabulary of 1000 words. Further data will probably lower this figure, that is, some children with vocabularies of 600 to 700 words may have attained the complete sentence, although we know that three children were still in the short sentence stage when possessing vocabularies of 800 to 1100 words.

Sentence Length with French and German Children

A few instances of average length of sentence of children of other nationalities are available. Descoeudres (1921:195) cites a 33-month

old French boy with a vocabulary of 639 words whose sample sentences show that he was in the short sentence stage. Fifty-seven consecutive sentences of a six-year-old French boy reported by Piaget (1923: 310–315) average 7.3 words. As for German children, the Sterns (1907: 61–62) give 40 sentences of their three-year-old daughter and these average 7 words; Schlag collected 11,270 sentences of a four-year-old girl which averaged 4.6 words, while the 12,078 sentences of a class of eight-year-old children averaged 7 words. Thus it would appear that the short sentence and complete sentence stages occur with French and German children as well as with those speaking English.

Examples of Different Stages

Early Sentences: H at 23 Months.—Diddie (lizard) way, way, way. Hocky (fox). Hooey (shoe). Daddy, look. Hingie, birdie, hingie (bird sits on my finger).

Short Sentences: H at 3 Years.—I not a bird. I tiny little baby. I date (great) big tild. I bigger Howard.

N at 5 Years.—I up a 'ee (tree). Bear 'an't get me. It gettin' dark? I 'ink 'o (think so). I did ha' two bir'day 'ake.

Complete Sentences: E at 33 Months.—Who is makin' dat little noise? Dere's a button dat's comin to see dem. Dat one is a kind of funny little dish, Mamma. I guess faiwies will eat out of dat dish.

D at 3½ Years.—If a house is on fire, put water on it. Then it will not make holes in it. And you can live in it forever.

Some Conclusions

The whole matter of the development of the sentence has been much neglected by American and English observers. We need data on the first appearance of the negative sentence, the first question, and the different kinds of questions—the whats, wheres, whens, hows and whys. Also the first use of compound and complex sentences should be noted. An excellent method of studying a child's speech development would be the recording of a half hour's or hour's conversation each month until the complete sentence was thoroughly established.

Collecting the vocabulary of any but a very young child is a laborious process; it means constant attendance on the child for three or four weeks and after that a great amount of work to assemble the results. The determination of a child's status in regard to his skill in

the use of the sentence is, on the contrary, simplicity itself—merely recording a series of 30 to 50 or more consecutive sentences and the necessary arithmetical computation. Let me emphasize that it is the *average* length of a series of sentences with which we are concerned (preferably ordinary conversation) and not the exceptional sentence. Moreover, although as few as 30 such sentences ought to show clearly a child's stage of speech development, for any comparative or detailed study there should be at least 100 sentences and preferably more. This average sentence length may well prove to be the most important single criterion for judging a child's progress in the attainment of adult language.

References

Boyd, Wm.: The Beginnings of Syntactial Speech. *Child-study*, Vol. VI, 1913, pp. 21–24, 47–51.

————: The Development of a Child's Vocabulary. *Ped. Sem.*, Vol. XXI, 1914, pp. 95–124.

Brandenburg, G. C.: The Language of a Three-year-old Child. *Ped. Sem.*, Vol. XXII, 1915, pp. 89–120.

———— and J.: Language Development during the Fourth Year. *Ped. Sem.*, Vol. XXIII, 1916, pp. 14–29.

————: Language Development during the Fourth Year. The Conversation. *Ped. Sem.*, Vol. XXVI, 1919, pp. 27–40.

Descoeudres, A.: "Le Développement de l'Enfant de Deux à Sept Ans." Neuchâtel, Delachaux, 1921, pp. 327.

Kirkpatrick, E. A.: "Fundamentals of Child Study." N. Y., Macmillan, 1910, pp. 384.

Miller, G. F. and M. D. and Nice, M. M.: A Boy's Vocabulary at Eighteen Months. *Proceedings* Oklahoma Academy Science, Vol. III, 1923, pp. 140–144.

Moore, K. C.: The Mental Development of a Child. *Psychology Review Monograph Supplement*, Vol. I, 1896, pp. 1–150.

Nice, M. M.: The Speech Development of a Child from Eighteen Months to Six Years. *Ped. Sem.*, Vol. XXIV, 1917, pp. 204–243.

————: Concerning All Day Conversations. *Ped. Sem.*, Vol. XXVII, 1920, pp. 166–177.

————: A Child Who Would Not Talk. *Ped. Sem.*, Vol. XXXII, 1925, pp. 105–143.

Pelsma, J.: A Child's Vocabulary and Its Development. *Ped. Sem.*, Vol. XVII, 1910, pp. 328–369.

Piaget, J.: "Le Langage et la Pensée chez l'Enfant." Neuchâtel, Delachaux, 1923, pp. 315.

Rowe, E. C & H. N. The Vocabulary of a Child at Four and Six Years of Age. *Ped. Sem.* Vol. X, 1913. pp. 187–208.

Schlag, J. Grundsätzliches zu einem Haüfigskeitwörterbuche der Kindersprache. *Ztschft. Päd. Psych.*, Vol. XVIII, 1917, pp. 216–226.

Scupin, E. and G. "Bubi's erste Kindheit." Leipzig, Grieben, 1907, pp, 264.
———: "Bubi im viertem bis sechsten Lebensjahre." Leipzig, Grieben, 1910, pp. 272.
Stern, C. and W.: Die Kindersprache. Leipzig, Barth., 1907, 2nd Ed., 1920, pp. 430.
Trettien, A. W.: Psychology of the Language Interest of Children. *Ped. Sem.*, Vol. XI, 1904, pp. 113–177.

THE QUANTITATIVE MEASUREMENT OF CERTAIN ASPECTS OF SOCIO-ECONOMIC STATUS

J. CROSBY CHAPMAN AND V. M. SIMS

Department of Education, Yale University

Those fortunate individuals who have convinced themselves that the so-called intelligence tests are measuring native intelligence must presumably at times have disconcerting moments of skepticism as to the degree to which their instruments are independent of home surroundings. Similarly those who are most convinced that the intelligence tests, especially of very young children, measure to an unfortunate degree favorable home opportunity have their more peaceful moments of belief. For, at times, they persuade themselves that the intelligence instruments are, after all is said to the contrary, extremely reliable even when employed on subjects coming from very diverse home conditions. In either case the factor of home conditions and its relationship to performance on tests and other school problems is, like a poor relative, ever with us.

There is no need to justify any study that has as its objective the construction of an instrument that attempts to measure the vague complex which, though analyzed later in this paper, may at the present moment be roughly defined as socio-economic status. Only as a quantitative instrument of this kind is available can we correct our reckonings within the field of ability measurement and subject the principle of equal exposure to scrutiny.

Several workers, including Van Denberg,[1] Holley,[2] Kornhauser,[3] and Counts,[4] have been interested in the question, though with reference to the importance of the subject socio-economic status has been relatively neglected. Experimenters within this field have been satisfied with isolated factors such as education of parents, occupation of parents, income, rentals, magazines and newspapers, libraries, clubs and organizations to which the parents belong. Holley concludes that the best single measure of home conditions is the number of books in the home library; Kornhauser comes to the conclusion that the

---

[1] "Elimination of Students in Public Secondary Schools." Teachers College.

[2] "Relationships between Persistence in School and Home Conditions. 15th year Book." Society for Study of Education, 1916.

[3] "Selective Character of American Secondary Education." U. of Chicago.

[4] Economic Standing of Parents and Intelligence of Children. *Journal of Educational Psychology*, Vol. IX, p. 159.

possession of a telephone is the most satisfactory measure; while Counts uses occupation of parents and telephone. In spite of careful search we have been unable to trace a single attempt to make a quantitative scale for measuring this complex. Though J. H. Williams[1] attempts to set up a rough descriptive scale for grading home conditions as observed by a social worker.

The fundamental weakness of all such studies using isolated factors as indices is that there is no possibility of quantitatively combining the findings. There is no method of summating the various factors that are assumed to measure social status. Each of the items may be symptomatic, but there is no indication, except general inference, as to the degree to which each is significant. It is as though one took a single intelligent act as symptomatic of the complex general intelligence. In measuring socio-economic status we must expect to have to use the same procedure as is employed in the measurement of intelligence. That is, we must measure different aspects of the total complex, decide in some manner the validity and intercorrelation of these aspects, and weight accordingly. Furthermore we must determine reliability, and eventually must submit the instrument to an external criterion as a test of its validity.

Unfortunately in the present study the approach was indirect. The data that we have been compelled to use were collected for the purpose of determining what factors influenced participation in extra-curricular activities. In the process of this investigation, however, the possibility of the construction of such an instrument as we have indicated became apparent. This second objective eventually eclipsed the previous purpose. In the questionnaire which had been sent out in connection with the first investigation of extra-curricular activities, 16 questions were pertinent to socio-economic status. These were as follows:

1. Do you have a telephone?
2. Do you work at some regular job out of school hours?
3. How many years did your father go to school?
4. How many years did your mother go to school?
5. Does your mother have a job outside of the home?
6. What language other than English is used in your home?
7. Is your home heated with a furnace?
8. Does your father own an auto which is not a truck?

---

[1] A Guide to the Grading of Homes. Whittier State School *Bulletin*, No. 7.

9. Have you a piano in your home?
10. Have you a victrola in your home?
11. How many rooms are there in your house? (Eliminated due to apartment house situation.)
12. How many magazines do you regularly get in your home?
13. How many books in your home?
14. What is your father's occupation? (Eliminated due to misunderstanding.)
15. Do your parents own their home? (Eliminated due to apartment house situation.)
16. How many daily papers do you regularly get in your home?

We cannot too clearly point out that the data collected in the questionnaire were not the most suitable for measuring socio-economic status. Had this study in its inception been framed with our final objective in mind, the questions employed would have been much more numerous and in certain cases very different from those that at the present time have to be employed. No claim therfore is made that these factors are the most significant for the measurement of socio-economic status. In fact, from later work we know that they are not. It is most unfortunate that the data on occupation of father due to a misunderstanding could not be employed in the scale. Nevertheless this material may be used to illustrate a procedure and it affords a reasonably useful instrument that can be employed while the tedious work of constructing better instruments is in progress. The authors are already engaged in constructing more elaborate instruments for use in various sociological settings.

As in the naming of all psychological and educational instruments, one always faces "the fallacy of labelling;" especially is this the case in the present instance when the label is such a loose term as socio-economic status. If pressed as to what is meant by socio-economic status one is compelled to answer that it is whatever this instrument measures. This, however, is a useless subterfuge. The need for some label which is generally interpretable is urgent. By socio-economic status we mean nothing more than the possession or non-possession of traits such as those above mentioned. If these traits are present, presumably the child has, both from the cultural and economic standpoint, a more favorable environment than is the case when these characteristics are absent. Furthermore it is reasonable to suppose that the more characteristics existing, the more favorable the environment. The whole study falls to the ground unless this assumption is allowed.

In the interests of the initial study of extra-curricular activities an elaborate questionnaire, including the selected socio-economic factors together with a large number of other questions with reference to age, scholarship, and extra-curricular activities had been given to the complete population of the New Haven academic high school. It is wholly upon these data that the scale is constructed. This endemic validity of the scale is unquestionably a weakness, but it is a weakness which is now being obviated.

It will be noticed that most of the questions with reference to socio-economic status are of the "yes or no" type of answer. In the interest of uniformity it was thought advisable to reduce all to this form. The data on such questions as extent of library and the amount of schooling of parents were examined from this point and at suitable places, such for example as possessing 125 books, or having an education beyond grade school, critical points were chosen. A library of the critical number of 125 or more books was regarded as a possession of the trait; likewise schooling on the part of parent beyond elementary school was regarded as possession of this trait. Similarly two papers or more, and two magazines or more constituted possession of the two characteristics, "papers," and "magazines." By this means all of the characteristics were reduced to the "all or none" principle. This has the advantage of affording uniformity of procedure in statistical analysis.

Columns 3 and 4 of Table I present the percentage possessing and not possessing each characteristic. Assuming the accuracy of these data and assuming that the distribution for each characteristic follows the over-worked Gaussian curve, we are able to assign a certain degree of socio-economic status in terms of sigma to the possession or non-possession of each trait. Of course it is well known that this vague complex which is termed socio-economic status is probably not distributed according to a normal distribution. Everything seems to suggest marked skewness, since low socio-economic status predominates. However in the absence of any definite distribution to follow other than the normal distribution, we make this assumption. It should be noted that having made this assumption, there is no object in working with extraordinary accuracy at any other point of the investigation. In educational statistics there is great danger of straining at gnats and swallowing camels—a common procedure which is not only bad scientific technique but also an enormous waste of time of the investigator.

By means of the Normal Probability Integral Table, the values for each trait were expressed in terms of sigma deviations from the average of the group. These are shown in columns 5 and 6. This introduced plus or minus deviations. To avoid the inconvenience of handling and interpreting negative values, it was decided to make all these values positive by adding 5 to each value. These results are shown in columns 7 and 8. Furthermore, since the inaccuracy of our assumptions does not justify great refinement, after multiplying by 10 we approximate these values in columns 9 and 10 to the first decimal place. Using these values for possession and non-possession of the traits, it is now possible to determine for each individual his total score on the combined set of characteristics. Only a scientific conscience long seared by contact with the addition of scores in intelligence tests and rating scales permits one to combine such diverse measures into a single scale. However this procedure needs no more apology than do hundreds of other procedures which have justified themselves in practice. The whole problem of the legitimacy of combining scores needs careful investigation in the light of the basic principles of scientific method.

TABLE I.—SHOWING SIGMA VALUES OF SOCIO-ECONOMIC STATUS, AND DERIVATIONS THEREFROM

(*P* = possession; *N-P* = non-possession of characteristic)

Characteristic	No. replies	Percentage		Sigma deviation		Sigma deviation +5		Values × ten and approximated	
		P	*N-P*	*P*	*N-P*	*P*	*N-P*	*P*	*N-P*
High school educated mother	2617	46	54	.86	− .73	5.86	4.27	59	43
Books (125+)	3018	41	59	.95	− .66	5.95	4.34	60	43
Piano	3153	78	22	.38	−1.35	5.22	3.65	52	37
Magazines (2+)	3103	67	33	.54	−1.10	5.54	3.90	55	39
Telephone	3153	63	37	.60	−1.02	5.60	3.98	56	40
Automobile	3153	36	64	1.02	− .60	6.02	4.40	50	44
High school educated father	2631	46	54	.86	− .73	5.86	4.27	59	43
Central heat	2956	71	29	.48	−1.18	5.48	3.82	55	38
English only language	3153	70	30	.50	−1.16	5.50	3.84	55	38
Victrola	3153	71	29	.48	−1.18	5.48	3.82	55	38
Daily papers (2+)	3072	53	43	.75	− .85	5.75	4.15	58	42
Mother not working	3153	94	6	.13	−1.99	5.13	3.01	51	30
Pupil not working	3153	74	26	.44	−1.25	5.44	3.75	54	38

At this point the procedure might stop and this scale might be regarded as final. But so far there has been no critical examination to

determine how satisfactory is the possession of each trait as a measure of socio-economic status. Apart from the manner in which the nature of distribution of traits in a high school population affects the values, it has been assumed implicitly that each trait is of equal importance in measuring socio-economic status.

Certainly all of the above characteristics are not of equal value as a measure of the complex "socio-economic status." In order to determine with rigidity the suitability of each item it would be necessary to set up some external criterion such as could be obtained by visitation of homes or by ranking of competent judges. Unfortunately no such external criterion is available. It is therefore necessary to resort to the somewhat dangerous principle of internal consistency. The two measures of internal consistency which are set up are: (1) the correlation of each characteristic with the total score on all the items with the exception of the one characteristic that is under examination; (2) the measurement of the degree of association between each of the separate characteristics.

In order to procure these measures of internal consistency without the unnecessary labor of working with the total population, 100 complete records were withdrawn in random order. That these 100 records taken from the four years of high school represent fairly adequately the total sample was shown by the degree to which the percentage of students possessing and not possessing the trait in the sampling followed closely the percentage for the total population.

In determining the degree of consistency between each separate item and the total on all characteristics irrespective of the particular item under examination, the Pearsonian Biserial r [1] was employed.

$$r = \frac{\frac{M_2 - M_1}{\sigma}}{\frac{z}{p} + \frac{z}{q}}$$

The values required for the use of the biserial coefficient are shown in Table II, and in the last column the correlations are reported. It will be seen that nine of the characteristics yield distinct positive correlations with the total measures. Two of them give somewhat low correlations, and two of them give practically no correlation. Nine, therefore, of these characteristics are reasonably consistent. Two of them although not so consistent still give some evidence of measuring a common factor. But the items "Mother not working"

[1] Kelley, T. L.: "Statistical Method," p. 249.

and "Pupil not working out of school hours" show no correlation and are discarded. The reason for the item "Pupil not working" giving such low correlation is due to the manner in which this factor is not equally incident on the sexes. No certain explanation can be offered for the characteristic, "Mother not working" giving such a low correlation. Possibly there are two conflicting factors; in the first place, unquestionably in many cases work on the part of the mother outside the home is indicative of low economic status, but in other cases it is not so, and here of course the added income of the mother increases the social status of the home.

TABLE II.—SHOWING VALUES ENTERING INTO THE BISERIAL COEFFICIENT OF CORRELATION

Characteristic	M_2	M_1	σ	$\frac{z}{p}+\frac{z}{q}$	Biserial r
High school educated mother	51.82	47.45	3.57	1.62	.67
Books (125+)	51.82	48.51	3.31	1.65	.62
Piano	50.50	46.74	3.56	1.69	.59
Magazines (2+)	50.33	46.98	3.56	1.63	.58
Telephone	50.47	47.34	3.58	1.62	.54
Automobile	51.53	48.33	3.73	1.64	.52
High school educated father	51.31	48.36	3.66	1.62	.50
Central heat	50.36	47.68	3.45	1.63	.48
English only language	50.45	47.56	3.65	1.62	.46
Victrola	50.15	47.94	3.78	1.64	.36
Daily papers (2+)	50.48	48.75	3.72	1.60	.28
Mother not working	49.59	48.07	3.95	2.06	.06
Pupil not working	49.39	49.38	3.92	1.69	.02

In the second method of determining internal consistency, Yule's coefficient of association[1] was employed to measure the intensity of

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}$$

association between the characteristics. In a study of this kind there was no justification in resorting to more accurate methods of measuring this degree of association. Table III presents these results and shows that with one or two exceptions the characteristics by no means measure precisely the same aspects of socio-economic status. The coefficients range from .93 which is the extreme positive deviate to

[1] Yule, G. U.: "Theory of Statistics," p. 38.

−.12, the extreme negative deviate. Twenty-three of these correlations are above .50 and only two of them are negative.

Had these items been more carefully selected and had the data been more universal we should have been justified in determining by partial or cumulative correlation an exact weighting for each characteristic in the light of these two measures of internal consistency. For the purpose of this study, however, we are justified in weighting approximately on the basis of the correlation of each characteristic with the total. Neglecting the degree of interrelation between the various separate measures, we are provided in column 1 of Table IV with the correlations quoted from Table II. Guided by the values of r we have weighted the factors as shown in the third column. In the following columns of the table there is shown the method of determining the final weightings which were employed in this test. In very case the measures were reduced to a positive basis by adding 5 to the sigma deviations and approximations were made to the first decimal place, and the decimal finally eliminated by multiplying by 10.

TABLE III.—SHOWING VALUES OF YULE'S COEFFICIENT OF ASSOCIATION

	High school education of mother	Books	Piano	Magazines	Telephone	Auto	High school education of father	Central heat	English alone spoken	Victrola	Daily papers	Average
High school education of mother	...	.72	.63	.85	.15	.32	.76	.63	.67	.12	.36	.52
Books	.72	.....	.75	.50	.38	.60	.57	.48	.72	.31	−.02	.50
Piano	.63	.75	...	.70	.57	.53	.26	.48	.42	.12	.24	.47
Magazines	.85	50	.70	...	.44	.35	.52	.53	.54	.50	.58	.55
Telephone	.15	.38	.57	.44	...	.93	.21	.43	.24	.39	.18	.40
Automobile	.32	.60	.53	.35	.93	...	.24	.62	.57	.48	.07	.47
High school education of father	.76	.57	.26	.52	.21	.24	...	.34	.49	.28	.24	.39
Central heat	.63	.48	.48	.53	.43	.62	.34	...	.34	.60	.40	.49
English alone spoken	.67	.72	.42	.54	.24	.57	.49	.34	.....	.28	−.12	.42
Victrola	.12	.31	.12	.50	.39	.48	.28	.60	.28	...	.11	.32
Daily papers	.36	−.02	.24	.58	.18	.07	.24	.40	−.12	.11	.....	.20

When the new values shown in the last column of Table IV are inserted for the 100 cases which have been selected at random, the distribution of scores shown in columns 1 and 2 of Table V results. This table gives us the idea of the significance of any particular score

TABLE IV.—SHOWING THE DERIVATION OF FINAL WEIGHTINGS FOR EACH CHARACTERISTIC

Characteristic	With total	Weighting	Original values of		Original values times weighting		5 added to columns 6 & 7		Final values × 10 approximated	
			P	*N-P*	*P*	*N-P*	*P*	*N-P*	*P*	*N-P*
			+	−	+	−				
High school education of mother	.67	10	.86	− .73	.86	− .73	5.86	4.27	59	43
Books (125+)	.62	10	.95	− .66	.95	− .66	5.95	4.34	60	43
Piano	.59	10	.38	−1.35	.38	−1.35	5.38	3.65	54	37
Magazines (2+)	.58	10	.54	−1.10	.54	−1.10	5.54	3.90	55	39
Telephone	.54	8	.60	−1.02	.48	−8.16	5.48	4.18	55	42
Auto	.52	8	1.02	− .60	.816	− .48	5.82	4.52	58	45
High school education of father	.50	8	.86	− .73	.688	−.584	5.69	4.42	57	44
Central heat	.48	8	.48	−1.18	.384	−.944	5.38	4.06	54	41
English alone spoken	.46	8	.50	−1.16	.40	−.928	5.40	4.07	54	41
Victrola	.36	6	.48	−1.18	.288	−.708	5.29	4.29	53	43
Daily papers (2+)	.28	4	.75	− .85	.30	−.34	5.30	4.66	53	47
Mother not working	.06	0	.13	−1.99	0	0	0	0		
Pupil not working	.02	0	.44	−1.25	0	0	0	0		

TABLE V.—SHOWING DISTRIBUTION OF SOCIO-ECONOMIC STATUS AND LEVELS DERIVED THEREFROM

Score	Frequency	Levels of socio-economic status		Percentiles	Score
55–56	6				
54–55	11	Indeterminate	9+	..	56+
53–etc.	3	Highest	9	95	55.1
52	6	Very high	8	88	54.5
51	8	High	7	79	52.8
50	14				
49	11	Medium high	6	66	51.0
48	7	Median	5	50	49.8
47	8	Medium low	4	34	48.1
46	6	Low	3	21	46.3
45	6				
44	11	Very low	2	12	44.9
43	1	Lowest	1	5	44.0
42	2	Indeterminate	1−	..	43−
	100				

and from it, for convenience is deduced levels of socio-economic status corresponding to the percentage indicated. It will be noticed that the distribution is not normal. At the higher extremes there is a tendency for the scale not to be discriminative. This is being rectified in the work which is now in progress.

Before using the scale in practical work, some idea of its degree of reliability is desirable. If we take a random sampling of the characteristics, six against five as shown in Groups I and II of Table VI, and correlate for these 100 cases the scores on the group of six characteristics as compared with the scores on the group of five character-

TABLE VI

Group I (Random)	Group II (Random)
Piano	High school education of mother
Magazines	Books
Telephone	Auto
High school education of father	Victrola
Central heat	Daily paper
English alone spoken	
Group A (Cultural)	Group B (Economic)
High school education of mother	Piano
Books	Telephone
Magazines	Auto
High school education of father	Central heat
Daily paper	English alone spoken
	Victrola

istics, an approximate idea of the reliability is given. The ordinary Pearsonian coefficient of correlation for these two halves of the score was .63. Another method of calculating the reliability is, instead of taking random halves of the scales, to take halves which stress on the one hand cultural factors and, on the other hand, economic factors (Groups A and B of Table VI). Under these conditions the correlation .49 was found. It is, of course, not strictly legitimate to apply the Brown coefficient of reliability[1] even to the random samplings, Groups I and II, in view of the fact that the elements comprising the two halves of the test are not, as the Brown procedure demands, interchangeable. However if it is employed merely to give us an approximate idea of the reliability of the whole scale, no great error is made. This coefficient of reliability is .77. It is totally ill egitimate to apply this formula to the selected groupings.

[1] Brown, W.: "Mental Measurements," p. 102.

This scale, in the form shown in Table VII, is now being employed for various purposes in connection with a study of the effect of home conditions, including and excluding language factor, on scores in intelligence tests. It is also part of a series of measures investigating the relationship between fecundity, intelligence and social status.

TABLE VII.—SOCIO-ECONOMIC SCALE

Characteristic	Answer	Weighting		Score given
		Possession	Non-possession	
Did your mother attend high school?.	.....	59	43	
About how many books are there in your home?.....	.....	60	43	
Do you have a piano in your home?...	.....	54	37	
How many magazines do you take regularly in your home?.....	.....	55	39	
Do you have a telephone in your home?	.....	55	42	
Do you have an auto other than a truck?.....	.....	58	45	
Did your father attend high school?..	.....	57	44	
Is your house heated by a big furnace in the basement?.....	.....	54	41	
Is English the only language spoken in your home?.....	.....	54	41	
Do you have a Victrola or Edison in your home?.....	.....	53	43	
How many daily papers do you regularly take in your home?.....	.....	53	47	

Books scored "possession" if 125 or more.
Magazines scored "possession" if 2 or more.
Papers scored "possession" if 2 or more.

GROUP TESTS AFTER SEVERAL YEARS

RUDOLF PINTNER

Teachers College, Columbia University

A study of group tests repeated in the same school raises two questions, the one practical and the other theoretical. The first question has to do with the educational and mental status of the school in spite of the shifting of the population due to the graduation of the older pupils and the arrival of many new pupils, particularly in the lower grades. The second or theoretical question has to do with the stability of the tests themselves and for this purpose we can consider only those children who were tested on both occasions.

Two schools for deaf children were given the Pintner Non-language and Educational Survey Tests on two occasions. School No. 1 was given the tests in 1919 and again in 1923, an interval of four years. School No. 2 was given the tests in 1919 and again in 1922, an interval of three years. The following table compares the percentage distribution of mental and educational indices in School No. 1 over a period of four years:

PERCENTAGE DISTRIBUTION

Indices	Mental indices		Educational indices	
	1919	1923	1919	1923
0– 19	1.4	1.2		
20– 39	14.5	16.6	13.1	4.9
40– 59	55.5	59.9	47.4	36.4
60– 79	27.7	22.3	36.4	49.4
80–100	0.7	0	2.9	9.3
n	137	162	137	162
Q_1	43	43	47	55
Median	52	52	56	63
Q_3	63.5	58	66	70
Q	10.25	7.5	9.5	7.5

The comparison between the mental indices in 1919 and 1923 shows that the mental make-up of the school in 1923 is practically the same as it was four years previous, although the population has

shifted decidedly. The turn-over in this school must be very great. Of the 162 pupils tested in 1923, only 26 were among those tested in 1919. In spite of this decided shift in the population, the intelligence of the pupils attending this school is evidently very constant. The comparison of the educational indices in 1919 and 1923 shows a decided rise in the four year period. The median index rises from 56 to 63. There is a distinct decrease of lower indices and an increase of higher ones, and this is in spite of the fact that the intelligence has remained the same. Evidently the school is making much better use of its mental resources in 1923 than it did in 1919. This factor can also be seen in a comparison of the differences between the mental and educational indices. The median difference in 1919 is +3 and this rises to +11 in 1923. Q_1 in 1919 is −4 and in 1923 it is zero. Q_3 similarly shifts from +13 to +21. In 1919, therefore, the school was accomplishing about what the ordinary deaf school accomplishes with children of like mentality. In 1923, the school is achieving much more than is generally achieved in the standard subjects taking into account the mental caliber of the pupils. In 1923 all except the lowest quarter of the pupils are accomplishing as much or more than is usually accomplished with children of like mentality. These facts would seem to point to better teaching, better supervision, or other similar factors during the four year interval, because the mentality of the pupils has remained constant.

With reference to School No. 2, a comparison of the mental and educational indices shows practically no change in either during the three-year interval. The median mental index moves from 55 to 58, while the median educational index remains at 54. In this school there is a very slight increase in the mentality of the pupils, but no change whatever in the educational attainments. Such a situation points probably to a routine way of teaching and supervising with no change over a period of three years.

The norms for our educational tests show, of course, merely what pupils are doing at present with present-day skill in teaching. We do not know what could be accomplished with more skillful instruction. If we notice that one school can accomplish more than present-day standards, we naturally begin to wonder why other schools cannot do likewise.

If we now turn to a study of those pupils who took the test on both occasions, we may learn something about the actual amount of improvement made and also something about the stability of the

test ratings. There were 26 pupils in School No. 1, who took the test in 1919 and again in 1923. The changes during the four year interval were as follows:

	1919	1923
Mean mental score	283	359
Equivalent MA (hearing)	11–3	13–1
Mean educational score	17	47
Equivalent EA (hearing)	8–4	11–2
Mean mental index (deaf)	54	54
Mean educational index (deaf)	56	63
Median difference	0	+10.5

The percentage of improvement in crude score in the mental test is 76 points or 27 per cent of the 1919 score. This is evidently a normal gain for deaf children, because we note that the mental index remains the same from 1919 to 1923. In the four year interval the growth in mental age for these deaf children is 1 year and 10 months in terms of growth for the average hearing child. This agrees with the general results from tests of deaf children indicating their retarded development. During this same interval, however, the growth in educational age in terms of hearing children is from an educational age of 8 years and 4 months, to 11 years and 2 months, that is, 2 years and 10 months. They have not grown educationally as fast as hearing children do, but this particular group has progressed educationally faster than mentally when contrasted with hearing children. This raises the general question as to whether the education of the deaf cannot be speeded up. This particular school seems to have accomplished something in this direction during a four year interval, and it would seem as if other schools might accomplish the same results. Another way of looking at the same problem is to compare the mental age in terms of hearing children with the educational age in terms of hearing children at the beginning and end of the four year interval. In 1919 the MA is 11–3 and the EA only 8–4, a difference of 2–11. In 1923 the MA is 13–1 and the EA is 11–2, and the difference has been reduced to 1–11. Or again we may raise the question as to why the deaf with a mentality equal to the hearing of 11–3 only accomplish educationally what the hearing of 8–4 accomplish? Undoubtedly the

difficulty in language encountered by the deaf is a potent factor. Is this the only factor? Are there not, perhaps, more efficient methods of teaching in general that might help to close this gap?

The Stability of the Ratings.—The stability of the mental and educational ratings can best be expressed in terms of correlation coefficients.

In School No. 1, with 26 cases over a four year interval, we have the following coefficients:

1. Mental score, 1919 with same in 1923, $r = .72$
2. Educational score, 1919 with same in 1923, $r = .70$
3. Mental index, 1919 with same in 1923, $r = .80$
4. Educational index, 1919 with same in 1923, $r = .80$

In School No. 2, for 91 pupils, we have:
Mental index, 1919 with same in 1922, $r = .75$
Educational index, 1919 with same in 1922, $r = .77$

All of these coefficients show a general tendency for the educational and mental ratings to remain fairly constant during a long interval of time. The coefficients obtained here are not as high as have been obtained for re-tests on the Stanford-Binet with hearing children. The correlation between the mental index in 1919 and the educational index in 1919 is +.66. The correlation between the mental index in 1919 and the educational index in 1923 is +.46. The value of the mental index for educational prediction has evidently decreased considerably over a period of four years.

In School No. 2 an attempt was made to obtain reports of positions and salaries of those pupils who had left the school since the tests were given in 1919. As usual the difficulty of follow-up work of this sort in a large city was tremendous. The wages of 23 boys and 16 girls were obtained. A rank correlation between wages and mental indices in 1919 resulted in a coefficient of −.10 for the boys and +.29 for the girls. The wages for both the boys and the girls ranged from $30 to $5 per week. The median mental index for 11 boys receiving $15 and over is 73, and for 12 boys receiving less than $15, the median mental index is 81, showing a slight tendency for the brighter boys to be receiving the lower wages. For the girls, the tendency is reversed for we have a median index of 67 for the 7 girls receiving $15 and over, and an index of 45 for the 9 girls receiving less than $15. These results are in general in accordance with other investigations which have shown that at least for the first few years after leaving school, there is

no relation between the wages received and the general intelligence of the child.

Conclusions.—In general, this study shows the desirability of investigating the results of tests repeated after long intervals, both from the point of view of the school itself and also from the point of view of the stability or reliability of the test ratings. This stability seems to be fairly great even for a short group test. In spite of the large pupil turn-over in the two schools, the constancy of the mental make-up of the children is marked. The marked increase in educational achievement in one school raises but does not answer some pertinent questions in the education of the deaf.

THE PRESENT STATUS OF THE INTRODUCTORY COURSE IN EDUCATIONAL PSYCHOLOGY IN AMERICAN INSTITUTIONS OF LEARNING[1]

O. B. DOUGLAS

Instructor in Educational Psychology, University of Texas

The purpose of this article is to indicate the present status of the introductory course in educational psychology in American universities and colleges. The data upon this problem were provided by sending a questionnaire to 73 educational institutions. This questionnaire, when properly answered, gave information on the five points listed below.

1. The college or university department in which educational psychology is offered;
2. The place of educational psychology as a factor in teacher-training;
3. The college year in which the introductory course of educational psychology is offered;
4. The subject-matter of educational psychology, and
5. The text and reference books now used in courses in educational psychology.

TABLE I.—SHOWING A LIST OF INSTITUTIONS OF LEARNING FROM WHICH REPLIES WERE RECEIVED

Baylor College, Belton, Texas.
Boston University, Boston, Mass.
Central Missouri State Teachers College, Warrensburg, Mo.
Clark University, Worcester, Mass.
College of the City of New York, New York, N. Y.
Colorado State Normal School, Gunnison, Colo.
Colorado State Teachers College, Greeley, Colo.
Columbia University, New York, N. Y.
De Pauw University, Greencastle, Ind.
Emory University, Atlanta, Ga.
George Peabody College for Teachers, Nashville, Tenn.
George Washington University, Washington, D. C.
Illinois State Normal University, Normal, Ill.
Indiana State Normal, Terre Haute, Ind.
Indiana University, Bloomington, Ind.
Kansas State Normal School, Emporia, Kan.

[1] The entire body of data upon which this article is based are to be found in a Master's thesis on "The History and Function of Laboratory Work in Educational Psychology," now in the library of the University of Texas. The writer is under obligations to Dr. C. T. Gray, University of Texas, for guidance in the preparation of this thesis, and for suggestions in the formulation of this article.

Leland Stanford Junior University, Stanford University, Cal.
Manchester College, North Manchester, Ind.
Montana State University, Missoula, Mont.
New York State College for Teachers, Albany, N. Y.
North Carolina College for Women, Greensboro, N. C.
Northeast Missouri State Teachers College, Kirksville, Mo.
Oberlin College, Oberlin, Ohio.
Ohio State University, Columbus, Ohio.
Ohio University, Athens, Ohio.
Pennsylvania State College, State College, Penn.
Purdue University, La Fayette, Ind.
Southwest Texas State Normal College, San Marcos, Texas.
State Normal College, Kent, Ohio.
State Normal School for Women, Farmville, Va.
State University of Iowa, Iowa City, Iowa.
Temple University, Philadelphia, Penn.
University of Alabama, University, Ala.
University of Arizona, Tuscon, Ariz.
University of Arkansas, Fayetteville, Ark.
University of California, Berkeley, Cal.
University of Chicago, Chicago, Ill.
University of Cincinnati, Cincinnati, Ohio.
University of Colorado, Boulder, Colo.
University of Dubuque, Dubuque, Iowa.
University of Florida, Gainsville, Fla.
University of Georgia, Athens, Ga.
University of Illinois, Urbana, Ill.
University of Kansas, Lawrence, Kan.
University of Kentucky, Lexington, Ky.
University of Michigan, Ann Arbor, Mich.
University of Minnesota, Minneapolis, Minn.
University of Missouri, Columbia, Mo.
University of Nebraska, Lincoln, Neb.
University of Oklahoma, Norman, Okla.
University of Pennsylvania, Philadelphia, Penna.
University of South Dakota, Vermilion, S. D.
University of Southern California, Los Angeles, Cal.
University of Tennessee, Knoxville, Tenn.
University of Texas, Austin, Texas.
University of Utah, Salt Lake City, Utah.
University of Virginia, Charlottesville, Va.
University of Washington, Seattle, Wash.
University of Wisconsin, Madison, Wis.
Vanderbilt University, Nashville, Tenn.
Washington University, St. Louis, Mo.
Western Illinois State Normal School, Macomb, Ill.
West Texas State Normal College, Canyon, Texas.
West Virginia University, Morgantown, W. Va.
Winthrop College, Rock Hill, S. C.

As can be seen, this is a representative list of American colleges and universities, and the total of 65 usable replies received gives a good basis for determining the tendencies which exist.

As indicated in the title the five points previously mentioned concern the first or introductory course in educational psychology and in the interpretation of the data which are presented this fact must be kept in mind.

In reply to the first question, "Are the courses in educational psychology given in the Department of Education or in the Department of Psychology?" 46 replies indicate that educational psychology is not given by the Department of Psychology, and 15 answers indicate that such a connection is now in existence. These data clearly indicate that the divorce of educational psychology from general psychology is not complete. It is possible that the increasing demand for teachers who have had training in terms of the situations which they will be called upon to meet has led to the courses under consideration being taken out of the hands of the general psychologist. Those who hold this point of view argue that the general psychologist is no more concerned with teaching than with any other vocation where his methods and principles were used, and that therefore teachers are best trained by those whose point of view is essentially pedagogic in its outlook.

On the other hand the feeling on the part of some psychologists that every phase of psychology should be under the same head, is evidently a strong factor in retarding the placing of educational psychology in the department of education. It is also possible that some of the institutions under consideration do not have a sufficient number of courses to warrant two departments.

Some significant information is contained in the replies received to question Number 2 of the questionnaire previously mentioned. This question is stated as follows; "Is the laboratory work done in the Department of Education, or in the Department of Psychology?" The 55 replies received are distributed as shown here.

Department of Education 35; Department of Psychology 20

In some instances the replies specifically stated that there were no laboratories in the department of education, and such experiments as were performed were done in the psychology laboratory. Also the fact was mentioned that, in some instances, only one laboratory instructor was provided for both departments, and therefore the laboratories were combined.

Another factor which has evidently helped to retard the final separation of the two types of laboratory work has been the lack of the contribution of definite and specific problems for the experiments in educational psychology in the form of laboratory manuals or text-books. Later results will show that there are no generally recognized and adopted books of this type. This may be due to the feeling that, after all, there is a very close connection between the experiment in educational psychology and the experiment in general psychology. Such a feeling would tend to modify any decisive steps toward bringing about a sharp separation of the two types of work. Another possible explanation may have to do with the difficulty which exists in devising experiments which grow out of distinct situations in the school room. There are those psychologists who place little value in any experiment for the prospective teacher which is not directly related to the problems of the school room.

The fact probably is, that since the experiment is a newer phase of the course in educational psychology than the subject-matter, it is, therefore, not as fully developed as other phases of the subject.

The replies to the question, "Is educational psychology a prerequisite to a teacher's certificate—to a degree in Education?" indicate, in a general way, the recognition and importance given educational psychology in the training of the teacher. In the 43 replies where educational psychology is required for a teacher's certificate, 33 States are represented, and in all of these, either the State Department of Education or the State Legislature, has to do with specifying the requirements for teachers' certificates. It may be concluded, therefore, that training in educational psychology is considered valuable for prospective teachers. Forty-five of the sixty-five institutions answering the questionnaire consider educational psychology of sufficient importance to require it for a degree in Education. In many of these institutions the courses offered are, from the catalogue descriptions, found to be very much alike, and especially is the similarity striking where these courses are offered in the same college year.

The fourth question of our questionnaire is stated as follows: "In what college year is educational psychology begun: Freshman, sophomore, junior, senior?" In the replies each school indicated the year in which such work was begun, and in cases of optional beginning, both years have been counted. According to this plan of computation the total number of replies is more than the number of schools replying, but our purpose of indicating the general tendency is clearly accomplished. The distribution is as follows:

Freshman	Sophomore	Junior	Senior
16	27	24	3

On the face of these divisions there seems to be no central tendency, and yet upon closer examination such a characteristic may be detected. The fact that the great majority of answers received place the work in either the sophomore or junior year, allows us to assume that there is at least a partial agreement concerning this point. Again, many States do not certificate teachers upon the completion of a single year of college work, and in these States the courses in educational psychology, although required for a teacher's certificate, are usually not offered until the beginning of the sophomore year. In those instances where educational psychology is offered in the first or freshman year, the State laws usually provide for allowing the certification of teachers on the basis of a single year of college work. Therefore, the apparent feeling that the course in educational psychology is the best single course for the prospective teacher, has caused this course to become a required course for those who plan to enter the profession of teaching.

The information contained in the answers to Question 5 may now be considered. The question reads: "Is educational psychology a prerequisite for all other courses in Education?" Number answering "yes," 46; number answering "no," 14. By briefly comparing the information contained in the answers to question 4 with that to be found in the answers to Question 5, it can be deduced that many colleges and universities do not offer courses in Education in their freshman, or first, year. It was seen in Question 4 that the general practice is to offer the first course in educational psychology in either the sophomore or junior year, and it is also true that, in the great majority of schools reporting, this first course is a prerequisite to all other courses in Education, it may be concluded, as before mentioned, that in the majority of institutions concerned, there are no freshman or first year courses offered in education.

Another fact which is more or less obvious on the face of the returns received, is that the elementary course in educational psychology forms, in the minds of the teachers of this subject, a satisfactory basis on which further educational training may be pursued. Its content, as offered in the institutions before mentioned, is evidently thought to be of such nature that it gives sufficient training in the fundamental principles of the educational process, and thereby obviates the necessity for taking further elementary work. It may be inferred, therefore, that there is coming to be more of a unanimity

of opinion concerning the content of this course in educational psychology, as well as a general agreement as to its relative value in terms of other courses in education.

The information necessary to answer the next question was obtained by requesting those answering the questionnaire to check from a list submitted those experiments which are performed in their educational psychology courses. These persons were also asked to add to the list any which are not included. In addition, they were requested to list the laboratory manuals used. A tabulation of the total number of times each experiment was checked is given below.

TABLE II.—SHOWING THE RANKINGS OF THE EXPERIMENTS SUBMITTED IN THE QUESTIONNAIRE

1.	Testing for individual differences	53
~~2.~~	~~Building up learning curves~~	~~48~~
2.	Building up learning curves	48
3.	Trial and error method of learning	43
4.	Testing for the intelligence quotient	43
5.	Immediate memory	42
6.	Memory; part *vs.* whole method	39
7.	Statistical method; distribution, correlation	39
8.	Illustrating factors of association	37
9.	Silent reading tests	35
10.	Arithmetic tests	33
11.	Laws of attention; intensity, contrast, etc	33
12.	Transfer of training	32
13.	Mirror drawing	31
14.	Finding plateaus in learning curves	29
15.	Tests for visual defects	28
16.	Influence of drill	25
17.	Experiments on testimony and report	25
18.	Introspection; vividness of imagination	24
19.	Experiments on sensation	24
20.	Testing for educational accomplishment	24
21.	Appreciption; hidden pictures, word completion	23
22.	Fluctuations of attention	22
23.	Tests for auditory defects	21
24.	Building up forgetting curves	21
25.	Testing for special abilities	15
26.	Substitution tests	2
27.	Spaced learning	2
28.	Relation of rate of learning to retentiveness	1
29.	Factors of space perception	1
30.	Sequence	1
31.	Suggestion	1

There seems to be a rather general agreement about the first 9 or 10 experiments on this list. It is interesting to note that in this first group of 10, including Number 7 there are 5 experiments from the field of tests, 2 concerned with the learning process, 2 in the field of memory, and 1 concerning the factors of association. This exceptional emphasis on tests in a laboratory course in educational psychology seems, at first appearance, to be out of place.

An explanation might be suggested at this point. It has been pointed out previously that there is an increasing demand for teachers who are trained in the application of objective methods to the problems of teaching. In a like manner the recent rapid growth of the field of tests and measurements has created a demand for teachers who, if not specifically trained in the manipulation of these devices, are at least in sympathy with their operation. The incorporation into the laboratory exercises of a few of these fundamental operations no doubt has served to promote a desire to cooperate in the administration of tests and measurements. This is especially true among those teachers who have had some training in this phase of educational psychology.

The emphasis on the learning process and memory is about equal, and this emphasis is continued in the second group of 10. In the second 10 experiments listed there is a noticeable absence of the test and measurement type. This is probably not due to any less consideration being given to this work, but to the fact that, because nearly all of the institutions checking this list give these test and measurement experiments, they are therefore included in the first 10 mentioned above. There appears to be, then, a greater unanimity of choice for this type of experiment than for any other particular type.

The distribution of the first 25 experiments of the list (which constitutes the original list as sent to the colleges and universities previously indicated) leads to the conclusion that there is some question as to just what experiments should be included in an introductory course of educational psychology. Also there seems to be little or no general agreement in the laboratory manuals which are available. In many instances the institutions provide their own laboratory manuals and, although the experiments in different colleges and universities may be alike in name, they usually vary to a greater or less degree in content.

Let us now notice the summary of the reports concerning the textbooks used in this course. In many instances several choices were given, and Table III has been made by taking each author the number of times he was named in the replies received.

TABLE III.—SHOWING THE NUMBER OF TIMES TEXTBOOKS BY THE FOLLOWING AUTHORS WERE MENTIONED AS BEING USED IN EDUCATIONAL PSYCHOLOGY COURSES

Starch	25
Gates	18
Strong	13
Pyle	12
Woodworth	12
Freeman	10
Seashore	9
Terman	9
Thorndike	9
Colvin	6
Judd	4
Averill	3
Parker	3
Gordon	2
Hollingsworth and Poffenberger	2
Hunter	2
James	2
Monroe	2
Titchener	2
Bennett	1
Bolton	1
Cameron	1
Emerson	1
Jennings	1
Langford	1
Charters	1
Meyers	1
Mead	1
Norsworth and Whitley	1
Proctor	1
Rugg	1
Watson	1

Many of the replies indicated the specific title of the book as well as the author, and this has made possible a rather interesting comparison. The list given on following page is composed of the first 10 authors named in Table II, together with the titles of their books and their respective dates of publication. It may be readily observed that those books which rank high in the list are new books.

TABLE IV.—SHOWING AUTHORS, TITLES, AND DATES OF PUBLICATION OF THE TEN TEXTBOOKS IN PSYCHOLOGY MOST OFTEN MENTIONED

Starch: Educational Psychology	1919
Gates: Psychology for Students of Education	1923
Strong: Introductory Psychology for Teachers	1922
Woodworth: Psychology	1921
Pyle: The Psychology of Learning	1921
Freeman: How Children Learn	1917
Thorndike: Psychology of Learning	1913
Colvin: The Learning Process	1911
Terman: The Measurement of Intelligence	1916
Seashore: Introduction to Psychology	1923

There are several significant indications with relation to this list which should be mentioned. The fact that many of the books are of very recent publication, and that they have been so generally accepted, leads to the conclusion that those whose business it is to chose textbooks in educational psychology do not seem to have been very well satisfied with books previously published. Again, the present tendency among teachers of educational psychology is to use books which have been written specifically for such courses. In view of this fact there will be a tendency to discard those books not so specifically written, and those discarded will include many books on general psychology which are now, or have very recently been, used in educational psychology courses.

There is also another factor which is worthy of consideration. Since the great majority of colleges and universities insist on giving this course in the sophomore or junior year, the prospective author has a rather definite level of mental activity and experience by which the subject-matter of his book may be gauged. This has probably led to a more uniform content in those books of recent publication which, doubtless, is one reason for the favorable reception given them.

In 1922, Remmers and Knight attempted to ascertain the text books which were then being used in educational psychology.[1]

They wrote to the leading colleges and universities of America and received a total of 41 usable replies. Since each reply indicated the books used both as texts and references, a summary of these answers gives a table of ranking much after the fashion of that given above. This ranking, together with the number of times each author was mentioned, is shown in Table V.

---

[1] *Journal of Educational Psychology*, Vol. XIII, Sept., 1922, pp. 399–407.

TABLE V.—SHOWING THE NUMBER OF TIMES TEXTBOOKS BY THE FOLLOWING AUTHORS WERE MENTIONED AS BEING USED IN EDUCATIONAL PSYCHOLOGY IN SEPTEMBER, 1922

Starch	11
Thorndike	10
Colvin	7
Seashore	6
Freeman	5
Pillsbury	5
Monroe	4
Woodworth	4
Averill	3
Terman	3
Angell	2
Bagley	2
Betts	2
Gordon	2
Kirkpatrick	2
La Rue	2
Norseworthy and Whitley	2
Strayer	2
Strong	2
Baldwin	1
Breese	1
Bolton	1
Calkins	1
Halleck	1
Hunter	1
Hollingsworth and Poffenberger	1
James	1
Judd	1
Ogden	1
Parker	1
Peterson	1
Tracy	1
Waddle	1
Watson	1

It is interesting, at this point, to indicate a comparison of the rankings given the various authors in September, 1922, and May, 1924. This is done in Table VI.

TABLE VI.—COMPARISON OF THE TEXT-BOOKS USED IN EDUCATIONAL PSYCHOLOGY IN SEPTEMBER 1922 AND MARCH 1924

Author	1924 rank	1922 rank
Starch	1	1
Gates	2	no rank
Strong	3	8
Woodworth	4	6
Pyle	4	no rank
Freeman	5	5
Thorndike	6	2
Terman	6	7
Seashore	6	4
Colvin	7	3
Judd	8	9
Averill	9	7
Parker	9	9
Gordon	10	8
Hollingsworth and Poffenberger	10	9
Hunter	10	9
James	10	9
Pillsbury	10	5
Titchener	10	no rank
Bolton	11	9

It should be explained that a few of the authors whose names appear in the 1924 group do not occur in the ranking of 1922. These have been marked "no rank." In most instances their names did not appear probably because the books referred to in the 1924 group had not been published when the 1922 ranking was made.

It will be observed that the lack of similarity between the two rankings can hardly be accounted for in terms of the time which intervened, since this amounted to only about 20 months. The fact that there has been such a rapid shifting of text and reference books is further proof of the lack of more than a very few books which have proven satisfactory. It seems that some of the newest books have proven most satisfactory, although there are some which rank high in both groups.

As has been previously mentioned, the above ranking was made from the number of times each author was mentioned. Since these replies were received from schools which begin their educational psychology courses anywhere from the freshmen to the senior year of college work,

the question naturally arises as to which books are used in each of the different years, and if there seems to be any definite agreement as to the particular books used in certain years of college work. The comparison, together with the number of preferences given for each book in each year, is given herewith.

TABLE VII.—SHOWING THE DISTRIBUTION AMONG THE FOUR COLLEGE YEARS OF THE DATA GIVEN IN TABLE V

Freshman		Sophomore	
Woodworth	6	Starch	9
Terman	5	Gates	8
Starch	4	Strong	6
Freeman	3	Pyle	6
Strong	3	Seashore	5
Gates	2	Colvin	3
Hollingsworth and Poffenberger	2	Freeman	3
Hunter	2	Thorndike	3
Pyle	2	Parker	2
Averill	2	Terman	2
Charters	1	Woodworth	2
Colvin	1	Averill	1
Gordon	1	Bennett	1
Judd	1	Mead	1
Monroe	1	Pillsbury	1
Thorndike	1	Gordon	1
		James	1
		Norseworthy and Whitley	1
		Monroe	1

Junior		Senior	
Starch	11	Gates	2
Gates	6	Starch	1
Strong	4	Woodworth	1
Pyle	4	Thorndike	1
Seashore	4	Terman	1
Freeman	4		
Thorndike	4		
Woodworth	3		
Judd	3		
Colvin	2		
Bolton	1		
Pillsbury	1		
James	1		
Watson	1		
Parker	1		
Terman	1		
Meyers	1		
Rugg	1		

The comparison of rankings shown above indicates a very close agreement between the text and reference books of educational psychology used in the sophomore and junior years, but these two differ rather widely from the ranking in the freshman year. The few schools beginning the work in the senior year mentioned only one book each and, except in two instances, these were different.

In concluding this article let us briefly summarize what has been said, and thereby indicate as specifically as possible what the present status of educational psychology is. It may be concluded that:

1. There is little or no uniformity of opinion as to the subject-matter of educational psychology except where the course is offered in a designated college year.

2. Consequently there appears to be no uniformly satisfactory textbooks for educational psychology.

3. There is a general, but not unanimous, agreement that educational psychology should be offered in the Department of Education.

4. That educational psychology should be a prerequisite to a teacher's certificate, as well as to a degree in Education, appears to be generally accepted.

5. The place for beginning courses of educational psychology seems to be in either the sophomore or junior year, while many schools offer the first course in the freshman year.

6. It appears to be generally accepted that a course in educational psychology should be required as a prerequisite for other courses in education.

7. The most satisfactory text and reference books used in educational psychology courses are very recent books. Starch, Gates, and Strong are the three favorite authors.

8. There has been a change in text and reference books during the last 20 months.

9. There is a decided agreement as to the proper books to be used where the first course of educational psychology is given in the sophomore or junior year.

THE PERMUTATION RELIABILITY COEFFICIENT

HERBERT A. TOOPS

Ohio State University

Particularly in subjective judgments of traits, as in rating scales, is a knowledge of the reliability coefficient of random judges a necessity. If we have two judges we desire to know to what extent they agree; accordingly we commonly compute the reliability coefficient of their judgments and this coefficient is a measure of the extent to which we may take a score of the one judge as the equivalent of the same score of the other judge. If there are three judges, we desire the *average* reliability coefficient computed by dividing $r_{12} + r_{13} + r_{23}$ by three. As an approximation we may plot the 3 paired series of measures X_1X_2, X_1X_3 and X_2X_3 on one plot and compute the value of the correlation coefficient from the scattergram thus resulting, calling the r thus secured the reliability coefficient of random judgment. To generalize, we may plot on one plot the *combinations* of paired measures of scores of the type X_1X_2, and obtain the correlation coefficient of the resulting plot.

For complete equivalence of judges, a given judgment, K, of Judge A should invariably be given for the same person by Judge B, by Judge C, etc., *irrespective of which judge happens first* to render his verdict, K, for the given person judged. This means that for complete equivalence, the means and standard deviations of judgments of different judges should be equal. If this is not approximately true, then the *permutations* of paired judgments when computed in one large correlation plot will yield a lower correlation coefficient than the *combinations* of judgments. The difference between the two coefficients will be an indication of whether or not the judges may be called equivalent judges; and consequently is a measure of whether or not the various judges' scores may be used without first reducing the scores to comparable measures, S_1's, by such formulæ as, $S_1 = \frac{X_1 - M_1}{\sigma_1} + K$; $S_2 = \frac{X_2 - M_2}{\sigma_2} + K$.

The combinations of judgments yield a correlation plot which may be solved for r by the usual formula,

$$R_{12} = \frac{\frac{N}{2}[\Sigma X^2 + \Sigma Y^2 - 2\Sigma(X - Y)^2] - \Sigma X \cdot \Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}} = \frac{\frac{N}{2}[A + B - C] - D \cdot E}{\sqrt{N \cdot A - D^2}\sqrt{N \cdot B - E^2}} \quad (1)$$

Fig. 1. Scattergram of judgments of 20 persons judged by Judges A and B. (Combinations of X,Y)

JUDGMENTS OF JUDGE B	JUDGMENTS OF JUDGE A: 1	2	3	4	5	6	7	8	9	10	Fr_Y
10								●		●	2
9									●		1
8						●					1
7							●				1
6					●	●		●		● P_3	4
5			● P_1		●						2
4				●	●		●				3
3		●	●								2
2	●	●		●							3
1	●										1
Fr_X	2	2	2	2	3	2	2	2	1	2	20

r_{AB} = .834 P = .830

Fig. 2. Scattergram of judgments of 20 persons judged by Judges A and B. (Permutations of judgments) Judgments X,Y are distinguished from Y,X judgments by plotting the former with solid circles, the latter with light circles.

	1	2	3	4	5	6	7	8	9	10	Fr_Y
10						P_4 ○		●		● ○	4
9									●○	•	2
8						●○				○	3
7				○			●○				3
6					●	●○		●○		● P_3	6
5			● P_1	○	●○	○					5
4		○		●○	●		●				5
3		●	●○		P_2 ○						4
2	●	●○	○	●							5
1	●○	○									3
Fr_X	3	5	4	5	5	6	3	3	2	4	40

The Pearson r of this diagram is found by solving P for Fig 1, which has only half the cases and so only half the plotting work. Pearson r for this diagram = .830

The permutations of judgments taken two at a time yield a correlation plot with twice as many cases in it as are involved in the above formula. The labor involved in plotting these cases can be saved by resort to the formula for the permutation reliability coefficient.

The problem is rendered more concrete by Figs. 1 and 2. Figure 1 gives the scattergram when judgments of Judge A are plotted against Judge B's judgments. In Fig. 2, each X, Y pair of the previous plot is entered on Fig. 1 again as light faced circles in the reversed order Y, X; thus the solid circle $P_1 = 3, 5$ is entered again in Fig. 2 as $P_2 = 5, 3$ and P_3 of Fig. 1 is entered a second time as P_4 on Fig. 2.

If we designate the permutation reliability coefficient by P, then the Pearson r of Fig. 2 is of the same numerical value as P of Fig. 1. An examination of the frequencies of Fig. 2 shows us that they are identical for X and Y thus simplifying the correlation formula greatly.[1] If in Formula 1 we substitute $2N$ for N, $\Sigma X + \Sigma Y$ for either ΣX or ΣY wherever occurring and $2\ \Sigma(X - Y)^2$ for $\Sigma(X - Y)^2$ of Formula 1, we obtain,

$$P = \frac{\frac{2N}{2}[\Sigma X^2 + \Sigma Y^2 + \Sigma Y^2 + \Sigma X^2 - 2\Sigma(X - Y)^2] - (\Sigma X + \Sigma Y)(\Sigma Y + \Sigma X)}{\sqrt{2N(\Sigma X^2 + \Sigma Y^2) - (\Sigma X + \Sigma Y)^2}\sqrt{2N(\Sigma Y^2 + \Sigma X^2) - (\Sigma Y + \Sigma X)^2}}$$

$$= \frac{2N[\Sigma X^2 + \Sigma Y^2 - \Sigma(X - Y)^2] - (\Sigma X - \Sigma Y)^2}{2N(\Sigma X^2 + \Sigma Y^2) - (\Sigma X + \Sigma Y)^2}$$

$$= \frac{[2N(\Sigma X^2 + \Sigma Y^2) - (\Sigma X + \Sigma Y)^2] - 2N\Sigma(X - Y)^2}{2N(\Sigma X^2 + \Sigma Y^2) - (\Sigma X + \Sigma Y)^2}$$

$$= 1 - \frac{2N\Sigma(X - Y)^2}{2N(\Sigma X^2 + \Sigma Y^2) - (\Sigma X + \Sigma Y)^2} \qquad (2)$$

On the Toops correlation form, this formula becomes, $r = 1 - \frac{2NC}{2N(A + B) - (D + E)^2}$. The formula is easy of solution in that there are no radical expressions involved.

This formula is applicable only to finding reliability coefficients of tests, judgments etc., where M_x and σ_x are assumed equal in all variables; *i.e.*, in finding the composite intercorrelation of X_1, X_2, X_3,

[1] This holds even though the solid dots may originally have been plotted from combinations of judgments from more than two judges. In either case, N of Formula 2 is the number of plot marks of Fig. 1, not necessarily the number of persons judged.

X_n, where all possible permutations of paired measures two at a time are used, thus assuming $M_{X_1} = M_{X_2} = M_{X_3}$; and $\sigma_{x_1} = \sigma_{x_2} = \sigma_{x_3}$, etc.

The above conditions may actually approximately hold, or they may be assumed to hold as an hypothesis for testing out the equivalency of judgment of the judges concerned.

TABLE I.—THE PERMUTATION RELIABILITY COEFFICIENT. METHOD OF SOLUTION
Its use in determining the reliability of rating of twelve different traits

Column	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*
Test	N	$\frac{\Sigma(X-Y)^2}{C}$	$= 2a \times b$	$\frac{\Sigma X^2}{A}$	$\frac{\Sigma Y^2}{B}$	$\frac{\Sigma X}{D}$	$\frac{\Sigma Y}{E}$	$=2N(A+B)$ $-(D+E)^2$	$P = 1 - \frac{C}{h}$	*Forward or Pearson r*	$= J - i$
a	30	31	1860	1438	1377	190	187	26771	.9305	.9324	.0019
b	30	23	1380	1817	1874	223	228	18059	.9236	.9279	.0043
c	60	86	10320	5510	5862	568	588	28304	.6354	.6657	.0303
d	35	20	1400	2735	2681	301	299	19120	.9268	.9296	.0028
e	48	24	2304	3683	3553	415	407	18972	.8786	.8855	.0069
f	60	75	9000	2815	2718	397	386	50871	.8231	.8325	.0094
g	110	112	24640	4406	3962	670	630	150960	.8368	.8572	.0204
h	120	130	31200	3461	3603	551	561	458816	.9320	.9327	.0007
i	120	156	37440	1969	1995	397	399	317744	.8822	.8822	.0000
j	120	306	73440	2528	2240	458	414	383936	.8087	.8180	.0093
k	36	124	8928	277	387	69	89	22844	.6092	.6395	.0303
l	24	63	3024	563	448	105	92	9719	.6889	.7194	.0305

For 12 separate traits the value of P is given in Column i of Table I, which shows the method of solution of Formula 2. Columns a, b, d, e, f, g, are empirical data taken from scattergrams of the relationship of Judge A's scores to Judge B's, similar to Fig. 1. The difference between r and P is given in Column k, and indicates high equivalency of judgment of all traits save for traits c, g, k, l. The method may be used to select, by elimination, equivalent judges for practical judging; it may also be used to advantage in such cases as testing whether the IQ's of a student in training correlate highly enough with IQ's of the instructor.[1] The reduction of the student's IQ's to standard measures is an unthinkable procedure in such a case. On the Stanford Scale it is not enough for instance that the student-tester's IQ's merely correlate highly with those of his instructor; he should also have the

[1] Herring, J. P.: Reliability of the Stanford and the Herring Revision of the Binet-Simon Scales. *Journal of Educational Psychology*, Vol. XV, No. 4, 1924, p. 222.

same average score and same variability of scores. P takes into account all these factors. Whenever $\frac{r}{r-P}$ becomes larger than some arbitrary value, say 60, or some other value empirically to be determined upon in the future, we may consider that the student has learned well enough how to use the scale and that further practice is unnecessary.

THE USE OF THE REGRESSION EQUATION IN DETERMINING THE APTITUDES OF AN INDIVIDUAL

CHAS. E. LIMP

Madison, Wis.

Before scientific vocational guidance can be made to function more effectively than it has in the past, it must be able to forecast not only abilities among individuals but the difference in abilities in the same individual. A great number of tests and test batteries have already been devised to predict ability or abilities of individuals, but to the knowledge of the writer there never has been a test or test battery devised to predict the difference of aptitudes in the same individual.

In a previous article, it was pointed out that perhaps the best differentiation of the four school subjects treated, was between English and typewriting where the highest k-value (Table IV) is .94 and the highest r (Table VI) is .495.[1] The best method of procedure for securing the difference between the above named aptitudes is perhaps by means of the regression equation. The equation serves not to predict the aptitude in English or typewriting but the *difference* between the two aptitudes, *i.e.*, it predicts in which aptitude an individual is likely to do better and by how much.

I

The criterion is derived from the school grades of 73 ninth year students in English and typewriting after these grades were transmuted into strictly comparable series having exactly the same mean and standard deviation.[2] Each student's grade in typewriting was subtracted from his English grade yielding a column of plus and minus differences. If the English grades exceed the typewriting grades, the scores are positive; if the typewriting grades exceed the English grades the scores are minus. The following will make this clear:

---

[1] Hull, Clark L. and Limp, Chas. E.: The Differentiation of the Aptitudes of an Individual by Means of Test Batteries. *Journal of Educational Psychology*, February, 1925, pp. 73–88.

[2] Hull, Clark L.: A Method of Converting Test Scores into Series Which Shall Have Any Assigned Mean and Degree of Dispersion. *Journal of Applied Psychology*, Vol. VI, 1922, pp. 298–300.

	School Grade English	School Grade Typewriting	Difference and Criterion
Subject No. 1	86	82	+ 4
Subject No. 2	78	90	−12
Subject No. 3	80	85	− 5
Subject No. 4	92	85	+ 7
Etc.			

TABLE I.—ZERO ORDER CORRELATIONS[1]

Test No. 3[2]	Best answers	(Terman)	+.398
Test No. 40	Word meaning	(Terman)	+.338
Test No. 4	Logical selection	(Terman)	+.155
Test No. 6	Sentence meaning	(Terman)	+.358
Test No. 7	Analogies	(Terman)	+.079
Test No. 8	Mixed sentences	(Terman)	+.247
Test No. 10	Number series	(Terman)	−.058
Test No. 13	Speed of writing	(Hoke)	+.185
Test No. 17	Spelling	(Hoke)	+.197
Test No. 18	Symbols	(Hoke)	−.037
Test No. 26	Volitional perseveration	(Downey)	+.074
Test No. 39	Multiplication	(Courtis)	+.022

[1] The numbering of the tests and the test scores were taken from Limp's original manuscript, referred to in the article (*op. cit.*).

[2] The 78 correlations (12 zero order correlations, Table I, and 66 intercorrelations Table II) were worked out with absolute accuracy in 16 continuous hours on the new correlation machine invented by Dr. Clark L. Hull, University of Wisconsin, a rate of 4⅝ correlations per hour. In order to have a check on the accuracy of the work, the writer had previously worked out 6 of the correlations using Crelle's and Barlow's Tables and the Monroe Computing Machine. Thirty-three hours were required, or an average of 5½ hours for each correlation as against 16 hours for 78 correlations with the machine.

The tests listed in Table I that yield a positive correlation with the criterion, indicate that a student will likely do better work in English than in typewriting, the size of the *r* indicating how much better. For instance Test Number 3 with a positive correlation of .398 indicates that a student will probably do much better work in English than in typewriting while Test Number 39 with a positive correlation of only .022 indicates that the student's abilities are practically equal in each

subject, the English being slightly favored. The opposite of this is true with those tests that yield a negative correlation with the criterion, *viz.*, Tests 10 and 18. The better a student does on a test, *i.e.*, the higher his score, the more likely he is to do better in English than in typewriting if the r is plus; if minus, the reverse is true.

II

In order that the R be as high as possible, the average error of prediction small, and the forecasts within a narrow range, the selection of tests for the battery is very significant. To be of greatest value, positive tests should correlate high with the criterion and low with each other; negative tests high with each other.

TABLE II.—INTERCORRELATION OF TESTS

	3	4	6	7	8	10	13	17	18	26	39	40
3	....	.452	.441	.174	.290	.276	.077	.135	−.113	−.172	.190	.475
4	.452	....	.371	.480	.451	.389	.243	.261	.137	.056	.377	.518
6	.441	.371	.....	.107	.352	.158	.099	.152	−.014	−.089	.103	.455
7	.174	.480	.107	....	.328	.378	.010	.254	.017	.087	.298	.303
8	.290	.451	.352	.328	....	.309	−.033	.238	.187	.026	.226	.293
10	.276	.389	.158	.378	.309	.....	.146	.159	−.055	.017	.270	.169
13	.077	.243	.099	.010	−.033	.146	.....	.201	.054	−.132	−.075	.015
17	.135	.261	.152	.254	.238	.159	.201	.....	.142	.208	.445	.043
18	−.113	.137	−.014	.017	.187	−.055	.054	.142	.....	.106	.277	−.047
26	−.172	.056	−.089	.087	.026	.017	−.132	−.208	.106	.....	−.121	−.025
39	.190	.377	.103	.298	.226	.270	−.075	.445	.277	−.121	.....	.127
40	.475	.518	.455	.303	.293	.169	.015	.043	−.047	−.025	.127	

$$PE = \frac{.6745\,(1 - r^2)}{\sqrt{N}}$$

N	±.0	±.1	±.2	±.3	±.4	±.5	±.6
73	.079	.078	.076	.072	.066	.059	.051

It is interesting to note that when the five positive tests (3, 6, 8, 17 and 40) that correlated highest with the criterion and lowest with each other were selected for the battery and the fourth order partials worked out, they yielded, respectively, .220, .145, .055, .136 and 126. Then Test 8, was dropped out of the battery and Test 10, which correlated negatively with the criterion to the extent of .058, was

substituted. This test not only raised the fourth partials of the other four tests—Test 3, from .220 to .264; Test 6, from .143 to .158; Test 17, from .136 to .185; Test 40 from .126 to .144—but yielded a negative fourth partial of .234 as against plus. 055 of test 8. Table III makes this clear.

TABLE III

Test	Zero order	4th partial	Test	Zero order	4th partial
3	+.398	+.220	3	+.398	+.264
6	+.358	+.143	6	+.358	+.158
8	+.247	+.055	10	−.058	−.234
17	+.197	+.136	17	+.197	+.185
40	+.338	+.126	40	+.338	+.144

It is also significant that by substituting Test 10 for Test 8 the R was raised from .481 to .518. This may seem like a small increase but it actually is a gain of almost 8 per cent in the differentiating efficiency of the test battery which is decidedly worth securing.

The regression equation

$$\overline{X}_c = +.490X_1 + .280X_2 - .321X_3 + .092X_4 + .171X_5 - 10.92,$$

(where the subscript c is the criterion and the subscripts 1, 2, 3, 4 and 5 are Tests 3, 6, 10, 17 and 40, respectively) gives the weighted scores for each test in the battery. By means of this equation the predictions or forecasts are made. In order to have a check on the R and prove the arithmetic of the work involved from the very beginning to the completion, the correlation between the criterion scores and the estimated scores was made. This gave an r of .518 which is a perfect check proving the accuracy of all the computations involved.

As stated in the beginning of this article this equation forecasts in which aptitude an individual is likely to do better and by how much. If the forecast is positive, it is likely that the student will do better work in English than in typewriting, the figure indicating the number of points. If the forecast is minus, the indication is that the person probably will do better work in typewriting than in English. If it is plus it indicates that a person probably will do better work in English than in typewriting. To make this clear, the scores made on each of the five tests in the battery, by subjects number 1, 2 and 3 are listed below. The last column gives the estimated criterion score.

	Test 3	Test 6	Test 10	Test 17	Test 40	Estimated criterion score
Subject No. 1.....	18	6	10	40	10	+1.8
Subject No. 2.....	8	3	14	16	0	−9.2
Subject No. 3.....	12	11	10	39	7	−1.5

In running these scores through the regression equation and performing the indicated arithmetic, subject Number 1 has an estimated score of +1.8; subject Number 2, −9.2 and subject Number 3, −1.5.

This means that student Number 1, is likely to do better work in English than in typewriting by 1.8 points of the ordinary marking scale and subjects Number 2 and 3 are likely to do better work in typewriting than in English by 9.2 and 1.5 points, respectively.

The average error of prediction of the 73 subjects, *i.e.*, the average of the differences between the criterion scores and the prediction scores, was found to be 6.8. Thus, from the above data it seems obvious that a regression equation can be constructed which serves to predict with a fair degree of accuracy the *difference* between aptitudes in an individual.

A MENTAL HYGIENE AND VOCATIONAL TEST

DONALD A. LAIRD

Colgate Psychological Laboratory, Hamilton, N. Y.

Tests of general mental ability, or alertness, or intelligence have been found to be of great usefulness in many fields. Yet their makers and users are clear spoken in stating that there is much about an individual that these tests leave unknown. This is not in any way a fault of intelligence tests—no one would expect a single type of test to tell all about an individual.

Some attempts have been made to fill in the lacunae in non-intelligence tests and a score of such tests could be noted.[1] I shall describe a test belonging in this group which experimental work indicates to be of considerable usefulness both in the way intelligence tests are useful and in other ways which some consider of even more importance.

The test might be descriptively called "A Test of Abnormal Mental Traits." It is based upon experience with the use of *Woodworth's Psychoneurotic Inventory* with students at Iowa University, the University of Wyoming, Yale College, and Washburn College (through the cooperation of Dr. Karl E. Menninger).

Woodworth's Inventory consists of a series of questions relating to psychoneurotic signs which are answered either *yes* or *no*. A question such as "Have you usually felt well and strong?" is answered either by yes or no. The modification which has been devised in this laboratory is to combine with these questions the type of response used in graphic rating scales. In addition many changes have been made in the questions. Those relating to events of childhood have been left out; others, while important, have been omitted because inquiries about defecation and urination in a mixed group have hindered the merited adoption of Woodworth's inventory. The questions have been grouped according to classical clinical entities, such as psychasthenoid, schizoid, neurasthenoid, hysteroid. The termination *oid* is used as suggesting a tendency in the direction of each of these clinical groups, rather than the actual presence of, for instance, neurasthenia.

Many questions have been added to give a completeness to the test, particularly in the case of introversion. One blank is used for introversion and another blank for the other questions.

---

[1] A summary of most of these is given by Symonds in *The Journal of Educational Psychology* for November, 1924.

The test is self-administering, and may be taken singly or in groups. It has been found good practice to have the subjects sit far enough apart that they are relieved of the possible embarrassment of their neighbors accidentally seeing their reports, which are intrinsically personal in nature.

The directions are:

This is not a contest. There is no time limit. Take all the time you need to describe yourself accurately.

You are asked questions about your actions and feelings which have important relations to your mental health and vocational aptitudes. You will answer these by making a check mark along the dotted line beside each question. Make a mark like this / extending through one of the dots.

One end of the dotted line means *little*, the other end *much;* in some questions it may mean *slightly* at one end and *strong* at the other. Each dot indicates a step from one extreme to the other.

Underneath the dotted line there are some phrases in small type. These are merely to help you get your bearings along the line. You do not have to make your check mark right over these, but may make it between them as well. In fact, about three-fourths of your check marks should be between these phrases.

You will proceed as follows: *First*, read the question in the large type. *Then* think how you have been in this respect *the last three months*. *Then* read the phrases in small type under the dotted line. Study these carefully, thinking about yourself the last three months as you do this. After you have reached a decision make the check mark at the place which best describes your actions or feelings.

Everything you place on these blanks will be given absolute confidence. The information is to help us guide you into your best life work and to help us guide you into this work with the best possible mental health. Nothing you check or write on the blanks will be held against you.

If you do not think that a check mark alone is the fairest way for you to answer some of the questions you may write in an explanation in addition to making a check, but be sure to check each line.

Remember: 1. Think how you have thought, felt, and acted the last three months.

2. Three out of every four checks should be between the phrases in small type.
3. Everything is confidential and to be used for your best interests.

The merits of a test such as this are: (1) That each person is described *directly* in terms of his actual behavior and not tricked indirectly into betraying his characteristics. (2) That each question is answered quantitatively, which is not possible with a yes-no type of response. (3) That traits which *exist*, rather than moral conceptions or "mental faculties," are reported on.

HAVE YOU HAD TROUBLE WALKING IN THE DARK?	never	only when dark and stumbled	occasionally	every time
HAVE YOU PREFERRED TO BE ALONE?	always sought company	enjoyed company but did not seek it	always sought solitude	
HOW HAVE YOU BEEN AT SELLING THINGS?	delighted in convincing customer	have sold many	will take orders, but not sell	avoided selling

Scoring and Norms

The scoring as a mental hygiene test is different from that as a vocational guidance test. The scoring as a mental hygiene test is accomplished by using a stencil which exposes the section of each dotted line in which a check mark is in the upper fourth of the distribution of check marks. Each check mark which is exposed by this stencil thus indicates that the individual deviates from normal in the trait. Norms are available for college students which show the distribution of the total deviations thus obtained in each division of the test.

Scoring as a vocational guidance test is accomplished by using a stencil which exposes the section of each dotted line in which fell the middle 50 per centum of check marks made by journalists, lawyers, salesmen, and so on. Data are being obtained from college graduates in different professions from which norms can be derived in cases of a significant difference.

Uses of the Test

Preliminary data indicate that sections of this are more indicative of failure in college courses than is an intelligence score. The test results have also proven useful in solving some knotty administrative problems in which mental orthosis was the prime need, and which would not have been realized through ordinary administrative channels.

Students deviating most from the normal can be spotted by the test and proper mental hygienic measures can be applied. As such a "spotting test" the blanks can be used to best advantage three times annually on each student; upon entering school in the fall, at the middle of the school year, and again just before the close of the year.

Thus it is possible to find those students in whom mental abnormalities are developing and to study the effects of various academic influences.

The research possibilities of the data obtained from this test are practically unlimited. Many important problems of abnormal psychology and college life can be brought nearer solution by these data, and our findings in this direction will be published after the work on reliability and validity is prepared.

At present the tests are being used at the following colleges: Wisconsin, Stanford, Illinois, Western Reserve, Skidmore, Buffalo, Colgate, Vassar, Olivet, Ohio University, Ohio State University, Purdue, and the University of Washington. In addition they are in use at one of the federal hospitals and are being adapted for use in Czecko-Slovakia. Several practising psychologists and psychiatrists are using the blanks.

Data on the reliability and validity of the test will be published shortly by Mr. Ralph K. Hoitsma. A comparison of college men and women will be published by Prof. Miriam C. Gould of Vassar, and Mr. T. McClumpha of Colgate. Statistics relating to the administrative uses and contributions to social and psychiatric problems will be ready later.

In the preparation of the present form assistance has been received from Prof. Miriam C. Gould, Dr. Harold E. Foster, and Dr. Karl E. Menninger.

AGREEMENT ON COOPERATION

C. SPEARMAN

University of London, University College

Once more, Professor Thorndike has shown that with his scientific eminence is consorted a no less admirable scientific sincerity. For a long time, his school and that to which I belong have been entangled in a controversy which has been sterile because the issues have been confused. Accordingly, I suggested to him that we should "agree as to what (if any) ascertained fact would be in harmony with the view of the one of us and not the other," so that any genuine difference between us might at last be submitted to a really crucial investigation.[1]

In response, Thorndike has now frankly and generously come forward with a definite scheme for such an investigation.[2] He has proposed the trying out of 8 tests of ability, as follows: Discrimination of pitch; memory of digits; three verbal tests (Opposites, Completion of Sentences, and Definitions); two arithmetical ones (Number Series and Problems); and Picture Completion. All these are to be applied to 10,000 16-year-olds, the accuracy being such as to secure reliability coefficients not less than 0.95. To this, I would only add that the subjects should be of the same sex, that they should have received reasonably similar education, and that all responses to one and the same test should be uniformly marked by the same person.

Now, on the terms of cooperation, Thorndike and I have to predict what appear to us respectively the most probable results. For my part, the following prophecy is here ventured. Take any four of the tests, but including only one at a time out of the three verbal ones, and only one of the two arithmetical ones. Let the picked four be denoted by the letters a, b, p, and q, whilst r as usual denotes the coefficient of correlation (product moment). Then, I anticipate, for every such four tests there will always be found valid the following equation,

$$r_{ap} \cdot r_{bq} = r_{bp} \cdot r_{aq},$$

---

[1] *Psychology Review*, 1922. *Journal of Educational Psychology*, Vol. XV, 1924, p. 393.

[2] *Journal of Educational Psychology*, Vol. XV, 1924, p. 304.

to the degree (and no more) that should be expected from the sampling "probable errors" involved.[1] As there are no less than 90 such equations at issue, and I am here standing for every single one of them, the prediction cannot be reproached with being meager. Yet more, I anticipate that the same equation will be fulfilled to a striking extent, though not so perfectly as in the other cases, when two or more of the verbal tests are included in the same group of four. The question will be a little further out, I add—but less so than might perhaps have been expected—if the two mathematical tests are admitted into one and the same group.

The turn passes now to Thorndike. Does he agree with this prediction, or does he not? Is he prepared to predict in any other respect? If so, I must be given a chance of agreeing or not. When this much has been settled, then—*if* our respective predictions do differ—we can set to work.

One thing more. Thorndike's proposal is rather heroic. 10,000 subjects! And no reliability less than .95! I would urge that we might profitably embark upon our enterprise in a tentative manner. We can begin with as little as 1000 subjects. Already with these, perhaps, the eventual issue will be clear enough to render the remaining 9000 dispensable. Even better and much safer is that as many independent investigators as possible should each try 100 subjects. Results can be averaged; mean deviations determined; anomalies detected, etc. As for the reliability coefficients, rigor seems to be on this occasion needless. For calculation at once shows that when (and only when) the above equation holds good of uncorrected coefficients, it also holds of the corrected ones.

---

[1] This degree of compensation required by the errors of sampling has recently been published by Professor Holzinger and myself (*British Journal of Psychology*, Vol. XV, 1924, pp. 17–19). Of course, even this only supplies at best an approximative value. With all experiments in every science whatever, the elimination of large disturbances only brings to light smaller ones, much as the departure of the sun makes visible the moon and stars. In the present case, we have especially to note that the above equation—in common with almost all current theory of correlations—assumes that the scores can be expressed with sufficient accuracy by a Taylorian expansion carried not further than the first differentials. Again, the calculation of the errors of sampling is, as usual, based upon assuming sufficiently "normal" frequency distributions. Furthermore there must at bottom exist innumerable complicating correlations; for instance, between memory of digits and the arithmetical performances. But all these disturbances appear to be so small, that in correlations up to 1000 subjects we may neglect them.

On contemplating this whole plan of collaboration between Thorndike, myself, and independent parties, I cannot refrain from a hope of its being widely adopted elsewhere, both in psychology and in other sciences. For surely in all alike, a most common occurrence is that the one contending school makes an elaborate research which it proclaims to support its view, whereat the other contendent objects that the research was so framed as to miss the real point. And so on in dreary interminable series. Could not, time and again, this squandering of precious scientific labor be escaped by an effective and sporting cooperation such as we are attempting here?

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY C. O. MATHEWS

Graduate Student, Teachers College, Columbia University

Intelligence Testing

Theory of Constancy of Intelligence. W. Stern. Translated by H. Kluver. The Psychological Clinic, March–April, 1925, 110–118. Results of the Stanford examinations are used to support a theoretical consideration of the constancy of mental development.

Group Differences in Intelligence Tests. Grace E. Manson. The Journal of Applied Psychology, June, 1925, 156–175. A study of the relative difficulty of the several types of questions for groups of different interests and educational status.

A Study of Four Primary Mental Tests. Ada B. Viele. The Elementary School Journal, May, 1925, 675–681. A comparative study using the Cole-Vincent, Dearborn, Detroit First Grade and Pintner-Cunningham Primary tests.

The Validity of Standards of Mental Measurement in Young Childhood. Helen T. Wooley. School and Society, April 18, 1925, 476–482. Retests of pre-school children show greater variability than those of older children. Some reasons for this suggested.

The Intellectual Resemblance of Twins. Eugene Shen. School and Society, May 16, 1925, 601–602. A critical discussion of Curtis Merriman's "The Intellectual Resemblance of Twins," Psychological Monographs, 1924.

Some Neglected Problems in Classifying High School Pupils. Richard D. Allen. Journal of Educational Research, May, 1925, 351–358. A plea for making use of educational intelligence tests for individual guidance as well as classification.

On the Provision of Alternative Forms of Examinations Equal in Difficulty. E. L. Thorndike. The Journal of Applied Psychology, Mar., 1925, 1–4. A note on the equivalence of twelve alternative forms of the Thorndike Intelligence Examination for High School Graduates.

Mental Tests for the Selection of University Students. Agnes L. Rogers. The British Journal of Psychology, April, 1925, 405–415. The Thorndike Examination can in 2½ hours forecast success with University work as accurately as High School records covering four years and college entrance examinations occupying 12 hours.

The Intelligence and Educational Achievement of Freshman Women of the University of Oklahoma. M. O. Wilson. School and Society, June 6, 1925, 693–694. Results on the Otis Advanced Intelligence Scale and their relation to age and educational marks. Comparisons are also made between rural and urban women.

The Relative Stability of Intelligence Levels in Colleges. Myra W. Kuenzel and Herbert A. Toops. School and Society, Aug. 8, 1925, 176–184. A comparison of the intelligence of freshmen in the various colleges of Ohio State University with the entire group, over a six year period.

Monthly Increments of Development in Infancy. Arnold Gesell. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1925, 203–208. Individual differences in the mental development of infants.

A Study of Performance Tests. Frances Gaw. The British Journal of Psychology, April, 1925, 374–392. An attempt to analyze the qualities measured by 14 well known performance tests and a standardization of these tests for English elementary school children.

Psychological Tests for College Freshmen. L. L. Thurstone. The Educational Record, April, 1925, 69–83. Data showing the comparative standing of 25 colleges on the same psychological tests.

Some Practical Results of Psychological Testing at Brown University. William R. Burwell and Andrew H. Mac Phail. School and Society, July 11, 48–56. A summary of findings after seven years of psychological testing and very careful follow up studies.

A Review of Racial Psychology. Thomas R. Garth. Psychological Bulletin, June, 1925, 343–364. A summary of contributions in this field from 1916 to 1924. All experiments tabulated; excellent bibliography.

A Study of Practical Ability. Margaret McFarlane. The British Journal of Psychology, Monograph supplements, VIII, 1925. Tests of the performance type were constructed to measure practical ability and the results correlated with general intelligence. A review of the literature on the nature of general intelligence is contained in the introduction.

Achievement Testing

A Study of the Achievement of High School Students in Chemistry. John W. Rauth. The Catholic Educational Review, May, 1925, 287–292. Some of the deficiencies in high school chemistry and their probable causes are pointed out.

Time Expressions Comprehended by Children of the Elementary School, II. Mary G. Kelty. The Elementary School Journal, May, 1925, 607–618. A study of children in the intermediate grades. Samples of tests for time-apprehension with results are given.

A Test for Classification of Students in Chemistry. Winifred F. Hyde and J. P. Guilford. Journal of Applied Psychology, June, 1925, 196–202. Data from the University of Nebraska showing the value of a psychological test for predicting success in Chemistry.

School Achievement and Success in Life. A. T. Poffenberger. The Journal of Applied Psychology, Mar., 1925, 22–28. The relation of class standing at West Point to success in civil and military life, to athletics and to dismissals from the army.

The Controlled vs. the Free Completion. Rose S. Malmud. The American Journal of Psychology, July, 1925, 401–411. The writer concludes from data presented that the controlled completion test more accurately measures language ability and is a finer index of general intelligence.

A Measure of Studiousness. Percival M. Symonds. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1925, 257–265. The difference between intelligence and achievement as measured by tests is suggested as a measure of studiousness.

Freshman Examination and Freshman Placement. Adam Leroy Jones. School and Society, April 11, 1925, 444–449. A preliminary report of the committee on College Personnel Technique of the Association of American Colleges. Discusses entrance and placement examinations.

An Experiment in Measuring Results of Fifth Grade Class Visits to an Art Museum. Katherine Gibson. School and Society, May 30, 1925, 658–662. Tests were given to 29 fifth grade classes and X, Y and Z groups were compared as to achievement in the test, as to environment of homes and as to retention.

Year to Year Differences in Spelling Ability Measured in Performance-group Grades. J. P. McCoy. School and Society, July 4, 1925, 23–24. Points out that "number of words right" is not a satisfactory measure of spelling ability.

Supervision through Simplified Testing. Charles E. Neville. The Elementary School Journal, May, 1925, 696–694. Tests in arithmetic, geography and history were constructed and grades compared as to their ability in the various processes.

Learning and Psychology of School Subjects

Experimental Investigations of Learning in the Case of Young Children. Arthur I. Gates. Journal of Educational Research, June, 1925, 41–48. Samples and suggested studies for research in learning in the case of children in the kindergarten and first two grades.

Needed Research in the Field of Learning. V. A. C. Henmon. Journal of Educational Research, May, 1925, 313–321. Numerous problems are suggested in all the fields of learning.

The Vocabulary of High School Pupils. F. P. O'Brien. Journal of Educational Research, May, 1925, 344–350. An analysis of the vocabulary used by 2542 pupils in junior and senior high schools.

A Suggestion for an Experiment on the Effect of Note-taking in Lecture Courses. Sadie Myers Shellow. The Journal of Applied Psychology, Mar., 1925, 69–74. A preliminary experiment discussed with suggestions for further experimentation.

A Comparative Study of Mathematical Ability of Boys and Girls in Secondary Schools. Annie E. Cameron. The British Journal of Psychology, July, 1925, 29–49. When mathematical ability is taken as a whole there appears to be no significant sex-difference.

Exceptional Children

Notwithstanding a High "IQ." Zoe I. Hirt. The Psychological Clinic, March–April, 1925, 128–132. The case study of a junior high school boy.

The Relationship between Suggestibility and Intelligence in Delinquents. John A. McGoech. The Psychological Clinic, March–April, 1925, 128–132. A study in which the relationship between suggestibility and intelligence is inverse.

A Psychological Study of a Group of Crippled Children of Various Types. Mable R. Fernald and Ada H. Arlitt, School and Society, April, 11, 1925. A study of the relationship of intelligence to disease and physical handicap.

X, Y and Z Pupils in the Detroit Public Schools. Harry J. Baker. School and Society, May 9, 1925, 570–572. Mental age levels and other quantitive differences between groups are given.

Fidelity of Report of Normal and Subnormal Children. John A. McGeoch. The American Journal of Psychology. July, 1925, 434–445. A comparison of the two groups as to range and accuracy of written and oral reports on object-card, picture and event.

Character and Personality

Analyzing Personality. George C. Brandenburg. The Journal of Applied Psychology, June, 1925, 139–155. Part I of a study attempting to show by ratings the relationships among various personality traits and their influence upon school and after-school life.

The Present Status of the Will-temperament Tests. Mark A. May. The Journal of Applied Psychology, Mar., 1925, 29–52. A critical analysis of significant data on the Downey Will-temperament Test and its Relation to character study.

A Critique of the Fables as a Moral Judgement Test. Gladys M. Lowe and Myra E. Shimberg. The Journal of Applied Psychology, Mar., 1925, 53–59. A comparison of the responses of 1443 delinquents with 445 non-delinquents on the Fables test in the Stanford Revision of the Binet-Simon scale.

Notes on Rating. Percival M. Symonds. The Journal of Applied Psychology, June, 1925, 188–195. Rating and ranking shown to be equally reliable. The amount of halo in rating is demonstrated by partial correlation.

Instincts and Emotions

What the Nursey Has to Say about Instincts. John B. Watson. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1925, 293–327. Results of experiments with infants and the behavioristic theory of instincts.

Experimental Studies of the Growth of the Emotions. John B. Watson. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1925, 328–348. Experiments with infants are summarized and the theory of emotions discussed.

Recent Experiments on How We Lose and Change our Emotional Equipment. John B. Watson. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1925, 349–371. Discussion of the methods used and results achieved.

A Study of the Emotions of Pre-school Children. Mary Cover Jones. School and Society, June 20, 1925, 755–758. Compares the use of different methods for curing fear.

Teachers' Marks

The Personal Element in Teachers' Marks. C. E. Hulten. Journal of Educational Research, June, 1925, 47–55. An experimental attempt to determine the usefulness of a corrective index for equating teachers' marks, concluding that it is impracticable since teachers are not consistent in giving high or low grades.

High School Marking Systems. C. W. Odell. The School Review, May, 1925, 346–354. Survey of the marking systems used in 281 Illinois High Schools.

Commutation of IQ's into Percentage Grades Corresponding to Those Commonly Used in Marking Scholarship. Gustave A. Feingold. Educational Administra-

tion and Supervision, April, 1925, 251–263. A suggestion to convert IQ's into terms comparable with the percentage grading system so as to compare capacity and achievement.

Comment on Mr. Feingold's Paper. Percival M. Symonds. Educational Administration and Supervision, April, 1925, 264–266. A criticism of the suggestion to use the percentage system and IQ as a basis for marking in high school.

Statistical Methods

The Validation of Statistical Procedure. S. A. Courtis. Journal of Educational Research, June, 1925, 31–40. A caution against using statistical procedures which are not appropriate to the situation.

A Printed Form for Computing the Standard Deviation on the Adding Machine. Herbert A. Toops. Journal of Educational Research, June, 1925, 56–58. A method which "eliminates the operation of making the extensions of the steps and their squares by frequencies of the respective steps."

Statistical Considerations in Interpreting the Effect of Training on Individual Differences. J. Crosby Chapman. Psychological Review, May, 1925, 224–234. It is impossible to measure this effect by the amount of time per unit of work since we have no absolute scale.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY JOHN HOCKETT
The Lincoln School of Teachers College

The Adolescent Girl, by Winifred Richmond. New York: The Macmillan Co., 1925. Pp. 212.

The author describes her work as "a book for parents and teachers." In the opinion of the reviewer, parents and teachers will find their needs served by the style and organization of the volume. As might be expected from the professional experience of the author, who is psychologist at the Government Hospital for the Insane, there is a decided emphasis upon psychological and social problems presented by unfortunate deviates. The abnormal girl and the delinquent girl occupy a third of the chapters. However, as parents and teachers are obliged in actual fact to spend a disproportionate amount of time and thought upon such individuals, the space devoted to them is justified.

Much of what is written here is applicable not only to the adolescent girl, but to adolescence regardless of sex. Nevertheless, there are problems of adolescence which concern the girl as girl. These arise from conflict between the reproductive interests and the self-assertive interests, which are natural antagonists, machinery and other forces in modern life having given freedom for the first time to the latter. Very sensibly the author restrains herself from offering a panacea. She points out that the data for solution can come only from experimental lives. After society has done its best to understand and educate her, the adolescent girl "must fight her own battles and learn her own lessons, as women have done since the world began."

LETA S. HOLLINGWORTH.

The Problem Child in School, by Mary B. Sayles, with a description of the purpose and scope of visiting teacher work, by Howard W. Nudd. New York Joint Committee on Methods of Preventing Delinquency, 1925. Pp. 287.

The case method of teaching, so long familiar in law and medicine, is now being brought over to some extent into educational psychology.

In recent years, several volumes have been made available, constituting a literature of instruction from this point of view. The book here considered, an outgrowth of study fostered by The Commonwealth Fund, is a worthy addition to this literature. In its pages are presented specific problems of habit formation, centering in parental attitudes, in feelings of inferiority, in dishonesty, in sex, and in miscellaneous situations, to which childhood is subject. The cases show how the visiting teacher may analyze the problem of the individual child, and may bring about cooperation between the home and the school. They also show how largely problems apparently of education are fundamentally problems of biology. Defects of the special senses, vital organs that function poorly, inferior intellects, and psychopathic temperaments are conspicuously frequent, both in the parents and in the children described. For instance, "The family backgrounds of the five children who stole vary greatly, yet one common weakness is to be found in them all. The only paternal figure which emerges is that of a step-father; one own father had died insane, after a youth spent in dissipation, two were deserters, a fourth bore a dubious reputation, the fifth was a positive menace." Again, "Steven and Alice, Jim and Herbert, each with his or her special degree and variety of dullness, stand for thousands of children who every year suffer, and are thwarted and warped, in a rigid school system constructed to fit the average child." Also, "By this time Mr. Lambert was again in jail, and Mrs. Lambert was almost frantic over the family situation" . . .

Obviously enough from these recitals, the visiting teacher deals with human material of the poorest quality. Occasionally she finds a basis for genuine constructive effort, giving promise of permanent effect, but much of her work is palliative only, under present conditions. Indirectly, however, such case work as that recorded here will lead to an understanding of the way in which effective education must base itself on human limitations, and hence to abandonment of present impossible ideals, which create "problems." Money spent in such studies is therefore money truly given to promote human welfare.

LETA S. HOLLINGWORTH.